

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2003, 08:18:04 ; Search time 17 Seconds
(without alignments)
365.149 Million cell updates/sec

Title: US-09-877-160A-1
Perfect score: 662
Sequence: 1 MKFIIAFFVATLAVMTVSGE.....GDILRLDKBEARVKKIEV 132

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	657	99.2	132	1 ALL5_DERPT	P14004 dermatophag
2	273.5	41.3	134	1 ALL5_BLOTA	O96870 blomia trop
3	187	28.2	110	1 ALL5_LEPDS	Q9u5p2 lepidoglyph
4	90	13.6	886	1 RA50_SULAC	O33600 sulfolobus
5	88	13.3	585	1 YHE0_YEAST	P38748 saccharomyc
6	87.5	13.2	2230	1 GOG4_HUMAN	Q13439 homo sapien
7	87	13.1	757	1 PIBF_HUMAN	Q8wxw3 homo sapien
8	86.5	13.1	978	1 RA50_AQUAE	O67124 aquifex aeo
9	85.5	12.9	254	1 PNM_DROME	Q9vtz6 drosophila
10	85	12.8	784	1 LON_ECOLI	P08177 escherichia
11	85	12.8	1433	1 REST_CHICK	O42184 gallus gall
12	84	12.7	166	1 YB38_MYCPN	P75260 mycoplasma
13	83.5	12.6	603	1 UVRB_BORBU	O51413 borrelia bu
14	83	12.5	2245	1 MYSJ_DICDI	P54697 dictyosteli
15	82.5	12.5	878	1 SYA_THETN	Q8rah4 thermoanaer
16	82.5	12.5	1978	1 MYHB_CHICK	P10587 gallus gall
17	82	12.4	144	1 Y047_ARCFU	O30189 archaeoglob
18	81.5	12.3	882	1 RA50_PRCFU	P58301 pyrococcus
19	81	12.2	886	1 CEP2_HUMAN	O29230 archaeoglob
20	80.5	12.2	2442	1 TOP1_BACSU	Q9bv73 homo sapien
21	80	12.1	691	1 LON_ERWAM	P39814 bacillus su
22	80	12.1	784	1 MYHB_HUMAN	P46067 erwinia amy
23	80	12.1	1972	1 REC7_HUMAN	P35749 homo sapien
24	79.5	12.0	565	1 MDP1_YEAST	Q9k974 bacillus ha
25	79.5	12.0	1875	1 RBP1_PLAVB	Q02455 saccharomyc
26	79.5	12.0	2869	1 PKB7_WHEAT	Q00798 plasmodium
27	79	11.9	559	1 EP45_XENLA	Q43207 triticum ae
28	78.5	11.9	436	1 MOEH_DROME	Q00387 xenopus lae
29	78.5	11.9	578	1 YNP9_CABEL	P46150 drosophila
30	78.5	11.9	705	1 YNP9_CABEL	P34562 caenorhabdi
31	78.5	11.9	845	1 SCP1_MESAU	Q60563 mesocricetu
32	78	11.8	879	1 RA50_SULTO	Q96yr5 sulfolobus
33	78	11.8	914	1 SYA_PYRAB	Q9uy36 pyrococcus

ALIGNMENTS

RESULT 1

ID	ALL5_DERPT	STANDARD;	PRT;	132 AA.
AC	P14004;			
DT	01-JAN-1990 (Rel. 13, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Mite allergen Der p 5 (Der p V) (IgE-binding allergen).			
GN	DERP5.			
OS	Dermatophagoides pteronyssinus (House-dust mite).			
OC	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;			
OC	Acariformes; Sarcoptiformes; Astigmata; Psoroptidia; Analgoidea;			
OC	Pyroglyphidae; Dermatophagoides.			
OX	NCBI_TaxID=6956;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=95096428; PubMed=7798547;			
RA	Lin K.L., Hsieh K.H., Thomas W.R., Chiang B.L., Chua K.Y.;			
RT	"Characterization of Der p V allergen, cDNA analysis, and IgE-mediated reactivity to the recombinant protein.";			
RT	J. Allergy Clin. Immunol. 94:989-996(1994).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90010801; PubMed=2794865;			
RA	Tovey E.R., Johnson M.C., Roche A.L., Cobon G.S., Baldo B.A.;			
RT	"Cloning and sequencing of a cDNA expressing a recombinant house dust mite protein that binds human IgE and corresponds to an important low molecular weight allergen.";			
RT	J. Exp. Med. 170:1457-1462(1989).			
RN	[3]			
RP	REVISION TO 132.			
RA	Cobon G.S.;			
RL	Submitted (FEB-1990) to the EMBL/GenBank/DBJ databases.			
CC	-!- SIMILARITY: BELONGS TO THE MITE GROUP 5 ALLERGEN FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; S76337; AAB32841.1; -			
DR	EMBL; S76340; AAB32842.1; -			
DR	EMBL; X17699; CAA35692.1; ALT_INIT.			
KW	Allergen.			
FT	VARIANT 61 61 E -> A.			
SQ	SEQUENCE 132 AA; 15615 MW; 3C840C3C26707C05 CRC64;			

Query Match 99.2%; Score 657; DB 1; Length 132;
Best Local Similarity 99.2%; Pred. No. 2.2e-43;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MKFIIAFFVATLAVMTVSGEDKKHYQNEFDLRLMERHEQIKKGELALFYLOEQINHEFE 60
|||||

Db 1 MKFIIAFFVATLAVMTVSGEDKKHDYQNEFFDLLEMERIHEQIKKGELALFYLQEQINHFE 60

QY 61 EKPTKEMKDKIVAEMDTIIAMIDGVRGVLDRMQKDLDFEYQYNLEMLKSGDILERDL 120

Db 61 EKPTKEMKDKIVAEMDTIIAMIDGVRGVLDRMQKDLDFEYQYNLEMAKSGDILERDL 120

QY 121 KKEEARVKKIEV 132

Db 121 KKEEARVKKIEV 132

RESULT 2

ALL5_BLOTA STANDARD; PRT; 134 AA.

AC Q96870; Q17283;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Mite allergen Blo t 5.

GN BLOT5.

OS Blomia tropicalis (Mite).

OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;

OC Acariformes; Sarcoptiformes; Astigmata; Glycyphagoidea;

OC Echimyopodidae; Blomia.

OX NCBI_TaxID=40697;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97154616; PubMed=9001334;

RA Arruda L.K., Vailes L.D., Platts-Mills T.A.E., Fernandez-Caldas E.,

RA Monteleagre F., Lin K.-L., Chua K.-Y., Rizzo M.C., Nasputz C.K.,

RA Chapman M.D.;

RT "Sensitization to Blomia tropicalis in patients with asthma and

RT identification of allergen Blo t 5.";

RL Am. J. Respir. Crit. Care Med. 155:343-350(1997).

RN [2]

RP SEQUENCE OF 63-134 FROM N.A.

RX MEDLINE=96426241; PubMed=8828535;

RA Caraballo L., Avjoglu A., Marrugo J., Puerta L., Marsh D.;

RT "Cloning and expression of complementary DNA coding for an allergen

RT with common antibody-binding specificities with three allergens of the

RT house dust mite Blomia tropicalis.";

RL J. Allergy Clin. Immunol. 98:573-579(1996).

CC -!- SIMILARITY: BELONGS TO THE MITE GROUP 5 ALLERGEN FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

DR EMBL; U59102; AAD10850.1; -.

KW Allergen.

FT CONFLICT 63 82 NKSKELOEKIIRELDVWCAM -> RAKELQRKSFENIMLF

FT APL (IN REF. 2).

FT Q -> P (IN REF. 2).

FT CONFLICT 126 126

FT SEQUENCE 134 AA; 15642 MW; F26D763475C8555FD CRC64;

Query Match 41.3%; Score 273.5; DB 1; Length 134;

Best Local Similarity 42.5%; Pred. No. 2.2e-14;

Matches 57; Conservative 32; Mismatches 40; Indels 5; Gaps 2;

QY 1 MKF--IIAFFVATLAVMTVSGEDKKHDYQNEFFDLLEMERIHEQIKKGELALFYLQEQIN 57

Db 1 MKFAIVLIACFAA--SVLAQEHKPKKDDFRNEFDHLLIEQANHAIEKGEHQLLYLQHQLD 58

QY 58 HFEKPTKEMKDKIVAEMDTIIAMIDGVRGVLDRMQKDLDFEYQYNLEMLKSGDILE 117

Db 59 ELNENKSKELOEKIIRELDVWCAMIEGAQGALERELKRTDLNILERFNYEBAQTLISKILL 118

QY 118 RDLKKEEARVKKIE 131

Db 119 KDLKETEQKVVDIQ 132

RESULT 3

ALL5_LEPDS STANDARD; PRT; 110 AA.

AC Q9U5P2;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Mite allergen Lep d 5 (Fragment).

OS Lepidoglyphus destructor (Storage mite).

OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;

OC Acariformes; Sarcoptiformes; Astigmata; Glycyphagoidea; Glycyphagidae;

OC Lepidoglyphus.

OX NCBI_TaxID=36936;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21099292; PubMed=11168362;

RA Eriksson T.L.J., Rasool O., Huecas S., Whitley P., Crameri R.,

RA Appenzeller U., Gafvelin G., van Hage-Hamsten M.;

RT "Cloning of three new allergens from the dust mite Lepidoglyphus

RT destructor using phage surface display technology.";

RL Eur. J. Biochem. 268:287-294(2001).

CC -!- SIMILARITY: BELONGS TO THE MITE GROUP 5 ALLERGEN FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

DR EMBL; AJ250278; CAB62212.1; -.

KW Allergen.

FT NON TER 1 1

FT SEQUENCE 110 AA; 12550 MW; D405050089E24CF1 CRC64;

Query Match 28.2%; Score 187; DB 1; Length 110;

Best Local Similarity 36.4%; Pred. No. 6.2e-08;

Matches 39; Conservative 24; Mismatches 44; Indels 0; Gaps 0;

QY 25 DYQNEFDLLMERIHEQIKKGELALFYLQEQINHFEKPTKEMKDKIVAEMDTIIAMIDG 84

Db 2 DFRNEFDRLIHMTTEQPAKLEQALAHLSHQVTELEKSKSKELKAQILREISIGLDFIDS 61

QY 85 VRGVLDRMQKDLDFEYQYNLEMLKSGDILERDLKKEEARVKKIE 131

Db 62 AKGHFERELKRALDLNLAENKFESALSTGAVLHKDLTALATKVKAE 108

RESULT 4

RASO SULAC STANDARD; PRT; 886 AA.

AC O33600;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE DNA double-strand break repair rad50 ATPase.

GN RAD50.

OS Sulfolobus acidocaldarius.

OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;

OC Sulfolobus.

OX NCBI_TaxID=2285;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 33909 / NCIB 11770 / DSM 639;

RX MEDLINE=97362314; PubMed=9211741;

RA Elie C., Baucher M.F., Fondrat C., Forterre P.;

RT "A protein related to eucaryal and bacterial DNA-motor proteins in the

RT hyperthermophilic archaeon Sulfolobus acidocaldarius.";

```

RL J. Mol. Evol. 45:107-114(1997).
CC -!- FUNCTION: Involved in DNA double-strand break repair (DSBR). The
CC rad50/mre11 complex possesses single-strand endonuclease activity
CC and ATP-dependent double-strand-specific exonuclease activity.
CC Rad50 provides an ATP-dependent control of mre11 by unwinding
CC and/or repositioning DNA ends into the mre11 active site (By
CC similarity).
CC -!- SUBUNIT: Forms a complex with mre11 (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y10687; CAA71688.1; -.
DR HAMAP; MF 00449; -.
DR InterPro; IPR003439; ABC transporter.
DR InterPro; IPR003395; SMC_N.
DR Pfam; PF04423; Rad50_zn_hook; 1.
DR Pfam; PF02463; SMC_N; 1.
DR ProDom; PD000006; ABC_transporter; 1.
KW DNA repair; Hydrolase; ATP-binding; Coiled coil.
FT NE_BIND 30 37 ATP (BY SIMILARITY).
FT DOMAIN 174 727 COILED COIL (POTENTIAL).
SQ SEQUENCE 886 AA; 103857 MW; 0390AE1403194104 CRC64;

Query Match 13.6%; Score 90; DB 1; Length 886;
Best Local Similarity 21.6%; Pred. No. 11;
Matches 32; Conservative 37; Mismatches 43; Indels 36; Gaps 6;

QY 20 EDKKHDYQNEFDLLMERIHEQIKKGELALF--YLQE---QINHFEKPTK----- 65
Db 223 EDELNQYNTFENRIKEIKVQVLDLISGELSVNKKIEIALRLKDFEEKEKYNKIETEVK 282

QY 66 ---EMKDKI--VAEMDTIIMIDGVRGVL-----DLMQRKLDLDFEQYNLEML 109
Db 283 ELDENREKINTISSFKSILVQIDSLKQINNVENDLKRKKEKLKRKKELEKEKEKQYEEIE 342

QY 110 KKSGLDILRLKKEEAR-----VKKIE 131
Db 343 KRKKELEKEKEKQYEEIEKRLTYVLKNIE 370

RESULT 5
YHBO_YEAST
ID YHBO_YEAST STANDARD; PRT; 585 AA.
AC P38748;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical 67.5 kDa protein in PRPS4-STE20 intergenic region.
GN YHL010C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8288c / AB972;
RX MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Favello A., Fulton L., Gattung S., Geisel C., Kirsten J.,
RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
RA Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
RA Nhan M., Rifkin L., Riles L., St Peter H., Trevasakis E., Vaughan K.,
RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
RA Vaudin M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VIII.";
```

```

Science 265:2077-2082(1994).
CC -!- SIMILARITY: TO C.ELEGANS EEED8.9 AND S.POMBE SPAC16E8.13.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U11582; AAB65064.1; -.
DR PIR; S46825; S46825.
DR SGD; S0001002; YHL010C.
DR InterPro; IPR001841; Znf_ring.
DR InterPro; IPR001607; Znf_UBP.
DR Pfam; PF00097; zf-C3HC4; 1.
DR Pfam; PF02148; zf-UBP; 1.
DR SMART; SM00184; RING; 1.
DR SMART; SM00290; Znf_UBP; 1.
DR PROSITE; PS00518; ZF_RING_1; FALSE_NEG.
DR PROSITE; PS0089; ZF_RING_2; 1.
KW Hypothetical protein; Zinc-finger.
FT ZN_FING 240 280 RING-TYPE.
SQ SEQUENCE 585 AA; 67503 MW; 88FF670CC73A1263 CRC64;

Query Match 13.3%; Score 88; DB 1; Length 585;
Best Local Similarity 23.7%; Pred. No. 11;
Matches 31; Conservative 23; Mismatches 47; Indels 30; Gaps 3;

QY 18 SGEDKKHDYQNEFDLLM-----ERIHEQIKKGELALFYLOEQINHEEKEPTKEMK 68
Db 376 SGDDDDNDNIGNSDELQNVVYGNRSKNGEKSNSKKDGEALANFLRHREYHLE----- 427

QY 69 DKIVAEIMTIAMIDGVRGVLDRMQRKLDLDFEQYNLEMLKKSGLDILRL----- 120
Db 428 -----YVQVLISQLESQREYVELKQEKDQTASDSSNVESLKKSMEDLKLQFOVTKEWQ 482

QY 121 KKEEARVKKIE 131
Db 483 KREMAQSKLE 493

RESULT 6
GOG4_HUMAN
ID GOG4_HUMAN STANDARD; PRT; 2230 AA.
AC Q13439; Q13270; Q13654; Q14436;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Golgi autoantigen, golgin subfamily A 4 (Trans-Golgi p230) (256 kDa
DE golgin) (Golgin-245) (72.1 protein).
GN GOLGA4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=96215236; PubMed=8626529;
RA Erlich R., Gleeson P.A., Campbell P., Dietzsch E., Toh B.-H.;
RT "Molecular characterization of trans-Golgi p230: a human peripheral
RT membrane protein encoded by a gene on chromosome 6p12-22 contains
RT extensive coiled-coil alpha-helical domains and a granin motif.";
RL J. Biol. Chem. 271:8328-8337(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Seelig H.P.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 131-2230 FROM N.A.
RC TISSUE=Placenta;
```



```
-----
CC EMBL; AF330046; AAL37481.1; -.
DR EMBL; Y09631; CAA70844.1; -.
DR EMBL; AL391384; CAC39624.1; -.
DR MIM; 607532; -.
FT CONFLICT 155 N -> S (IN REF. 2).
FT CONFLICT 333 R -> C (IN REF. 2).
FT CONFLICT 358 MYEKYVASR -> I (IN REF. 3).
FT CONFLICT 597 DLEHRKDQVTQLSPELDR -> RSGTSKGPSNTAFTRSLTE
      (IN REF. 2).
FT CONFLICT 610 610 P -> Q (IN REF. 3).
SQ SEQUENCE 757 AA; 89773 MW; B468EB2A071E5B8D CRC64;

Query Match      13.1%; Score 87; DB 1; Length 757;
Best Local Similarity 23.8%; Pred. No. 16;
Matches 29; Conservative 34; Mismatches 39; Indels 20; Gaps 5;

QY 12 LAVMTVSGEDKKHDYQNEFDLLMERIHEQIKKQELALFYLOEQINHFEEKPTKEMDKI 71
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 67 LSQKTMIDNLKVDYLTKE-ELEEKLNDAHFQKQLTLRLDNQL-AFQOKDASKYQELM 124

QY 72 VAEMDTIAMDGVRLDRLMQKDLDFEQVNLMLKSGDILE--RDLKKEEARVKK 129
   ||:|
Db 125 KQEMETI-----LLRQKQL---EETNLQREKAGDVRRNLRDPFELTEEQYIK 168

QY 130 IE I31
   :
Db 169 LK 170

RESULT 8
RA50 AQUAE
ID RA50_AQUAE STANDARD; PRT; 978 AA.
AC O67124;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable DNA double-strand break repair rad50 ATPase.
GN RAD50 OR AQ 1006.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RA "The complete genome of the hyperthermophilic bacterium Aquifex
RA aeolicus." ;
RA Nature 392:353-358(1998).
CC -!- FUNCTION: Involved in DNA double-strand break repair (DSBR). The
CC rad50/mrell complex possesses single-strand endonuclease activity
CC and ATP-dependent double-strand-specific exonuclease activity.
CC Rad50 provides an ATP-dependent control of mre11 by unwinding
CC and/or repositioning DNA ends into the mre11 active site (By
CC similarity).
CC -!- SUBUNIT: Forms a complex with mre11 (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; AE000718; AAC07092.1; -.
DR PIR; A70387; A70387.
DR HAMAP; MF_00449; -.
DR InterPro; IPR003439; ABC_transporter.
```

```
DR InterPro; IPR004592; SbcC.
DR InterPro; IPR002017; Spectrin.
DR Pfam; PF04430; DUF498; 1.
DR TIGRFAMS; TIGR00618; sbcc; 1.
KW DNA repair; Hydrolase; ATP-binding; Coiled coil; Complete proteome.
FT NP BIND 32 39 ATP (BY SIMILARITY).
FT DOMAIN 160 826 COILED COIL (POTENTIAL).
SQ SEQUENCE 978 AA; 115897 MW; 9B0F2BF51ADD1151 CRC64;

Query Match      13.1%; Score 86.5; DB 1; Length 978;
Best Local Similarity 24.4%; Pred. No. 23;
Matches 30; Conservative 33; Mismatches 41; Indels 19; Gaps 5;

QY 12 LAVMTVSGEDKKHDYQ---NEFDLLMERIHEQIKKQELALFYLOEQINHF-----EEK 62
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 355 LSQLSSSLKEKEREYEQAKQEF-----EDLSERVEKGGKILVAETEEKLEIKELFSEEEY 409

QY 63 PTKEMKDKIVAEMDTIAMDGVRLDRLMQR-KDLDFEQVNLMLKSGDILERDLK 121
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 410 TSLKMKERLLVELQRLKELKEKEGQLENLTQYKEKKKVKHEKVLNELKE----LERELK 465

QY 122 KEE 124
   :
Db 466 ERE 468

RESULT 9
PMM_DROME
ID PMM_DROME STANDARD; PRT; 254 AA.
AC Q9VTZ6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable phosphomannomutase (EC 5.4.2.8) (PMM).
GN CG10688.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.B., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
```

```
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -!- FUNCTION: INVOLVED IN THE SYNTHESIS OF THE GDP-MANNOSE AND
CC DOLICHOL-PHOSPHATE-MANNOSE REQUIRED FOR A NUMBER OF CRITICAL
CC MANNOSEYL TRANSFER REACTIONS (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: D-mannose 1-phosphate = D-mannose 6-phosphate.
CC -!- PATHWAY: Glycosylation; early steps of mannosylation.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE EUKARYOTIC PMM FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE003541; AAF49899.1; -.
DR FlyBase; FBgn0036300; CGI0688.
DR InterPro; IPR006379; HAD_SF_IIB.
DR InterPro; IPR005002; PMM.
DR Pfam; PF03332; PMM; 1.
DR TIGRFAMs; TIGR01484; HAD-SF-IIB; 1.
KW Hypothetical protein; Isomerase.
SQ SEQUENCE 254 AA; 29667 MW; 716E05E7BB2438BB CRC64;
Query Match 12.9%; Score 85.5; DB 1; Length 254;
Best Local Similarity 27.0%; Pred. No. 6.9;
Matches 37; Conservative 28; Mismatches 47; Indels 25; Gaps 7;
QY 10 ATLAVMTVSGEDKKHDYQ-----NEFDLIMERIHEQIKKGELALFYLQEQINHFEE 61
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::| 44
ATIGI--VGGSDLEKMFQELNGRKILNEFDIFPENGVLVIEGKGV--GKQNIIMHLGE 99
QY 62 KPTKEMKD---KIVAEMD-----TIIAMIDGVRGV--LDRLMQRKDLDFEYQYNLEMLK 110
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::| 100
ETVKRFNFVFLRYLSELDVPIKRGTFIEFRNGMVMVCPIGRQCTREERNMFAEYDIE--H 157
QY 111 KSGDILERDLKKEARV 127
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::| 158
KVREKMIKOLKQEFADV 174
RESULT 10
LON_ECOLI
ID LON_ECOLI STANDARD; PRT; 784 AA.
AC P08177; P78219;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ATP-dependent protease La (EC 3.4.21.53).
GN LON OR CAPR OR DEG OR MUC OR LOPA OR B0439 OR C0555.
OS Escherichia coli, and
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 217992;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=94124005; PubMed=8294008;
RA Thomas C.D., Modha J., Razzaq T.M., Cullis P.M., Rivett A.;
RT "Controlled high-level expression of the lon gene of Escherichia coli
RT allows overproduction of Lon protease.";
RL Gene 136:237-242(1993).
RN [2]
```

```
RP SEQUENCE FROM N.A.
RX MEDLINE=88251494; PubMed=3289547;
RA Amerik A.Y., Chistyakova L.G., Ostroumova N.I., Gurevich A.I.,
RA Antonov V.K.;
RT "Cloning, expression and structure of the functionally active
RT shortened lon gene in Escherichia coli.";
RL Bioorg. Khim. 14:408-411(1988).
RN [3]
RP SEQUENCE FROM N.A., AND MUTAGENESIS OF SER-679.
RX MEDLINE=94043004; PubMed=8226758;
RA Fischer H., Glockshuber R.;
RT "ATP hydrolysis is not stoichiometrically linked with proteolysis in
RT the ATP-dependent protease La from Escherichia coli.";
RL J. Biol. Chem. 268:22502-22507(1993).
RN [4]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 783-783.
RX MEDLINE=88298842; PubMed=3042779;
RA Chin D.T., Goff S.A., Webster T., Smith T., Goldberg A.L.;
RT "Sequence of the lon gene in Escherichia coli. A heat-shock gene
RT which encodes the ATP-dependent protease La.";
RL J. Biol. Chem. 263:11718-11728(1988).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [6]
RP SEQUENCE FROM N.A.
RA Roberts D., Allen E., Araujo R., Aparicio A., Chung E., Davis K.,
RA Duncan M., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
RA Lew H., Lin D., Namath A., Oefner P., Schramm S., Davis R.W.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=06:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
RN [8]
RP SEQUENCE OF 1-58 FROM N.A., AND SEQUENCE OF 1-14.
RX MEDLINE=85157418; PubMed=2984174;
RA Gayda R.C., Stephens P.E., Hewick R., Schoemaker J.M., Dreyer W.J.,
RA Markovitz A.;
RT "Regulatory region of the heat shock-inducible capR (lon) gene: DNA
RT and protein sequences.";
RL J. Bacteriol. 162:271-275(1985).
RN [9]
RP MUTAGENESIS OF LYS-362.
RX MEDLINE=95080406; PubMed=7988699;
RA Fischer H., Glockshuber R.;
RT "A point mutation within the ATP-binding sites inactivates both
RT catalytic functions of the ATP-dependent protease La (Lon) from
RT Escherichia coli.";
RL FEBS Lett. 356:101-103(1994).
RN [10]
RP MUTAGENESIS.
RX MEDLINE=99194734; PubMed=10094703;
RA Ebel W., Skinner M.M., Dierksen K.P., Scott J.M., Trempy J.E.;
RT "A conserved domain in Escherichia coli Lon protease is involved in
RT substrate discriminator activity.";
RL J. Bacteriol. 181:2236-2243(1999).
CC -!- FUNCTION: DEGRADES SHORT-LIVED REGULATORY AND ABNORMAL PROTEINS
CC IN PRESENCE OF ATP. DEGRADES THE REGULATORY PROTEINS RCSA AND
```


DR EMBL; AF045650; AAC03547.1; -;
DR EMBL; AF045651; AAC03548.1; -;
DR InterPro; IPR000938; CAP-Gly.
DR InterPro; IPR001878; Znf CCHC.
DR Pfam; PF01302; CAP_GLY; 2.
DR SMART; SM00343; Znf_C2HC; 1.
DR PROSITE; PS00845; CAP_GLY_1; 2.
DR PROSITE; PS0245; CAP_GLY_2; 2.
KW Cytoskeleton; Microtubules; Coiled coil; Repeat; Alternative splicing.
FT DOMAIN 79 121 CAP-GLY 1.
FT DOMAIN 144 207 SER-RICH.
FT DOMAIN 235 277 CAP-GLY 2.
FT DOMAIN 305 332 SER-RICH.
FT DOMAIN 351 1353 COILED COIL (POTENTIAL).
FT DOMAIN 1414 1427 CCHC-BOX.
FT VARSPPLIC 458 492 Missing (in isoform 2).
FT VARSPPLIC 458 492 /FTid=VSP_000761.
FT VARSPPLIC 458 492 TOTKLEHARIKELQSLFFKTKADKLQLEEDTR -> RK
FT VARSPPLIC 458 492 ROISEDPEN (in isoform 3).
FT VARSPPLIC 803 803 /FTid=VSP_000762.
FT VARSPPLIC 803 803 S -> GGSSKVS (in isoform 3).
FT VARSPPLIC 458 458 /FTid=VSP_000763.
FT VARSPPLIC 458 458 T -> RKQISEDPENT (in isoform 4).
FT CONFLICT 309 309 K -> R (IN REF. 2; AAC03547).
FT CONFLICT 440 440 E -> V (IN REF. 2; AAC03548).
SQ SEQUENCE 1433 AA; 161026 MW; 5631CE8683498E23 CRC64;

Query Match 12.8%; Score 85; DB 1; Length 1433;
Best Local Similarity 31.9%; Pred. No. 44;
Matches 36; Conservative 21; Mismatches 44; Indels 12; Gaps 5;

QY 22 KKHDYQNEFFLLMERIHEQIKKGELALFYLQEQINHFEEKPTKEMKDKIIVAEKMDTIAM 81
DB 1047 KKND-ELETAELKKQAEQAKADKRA-----EVLQTMKVKTK-KDAIHQEKIETLAS 1099

QY 82 IDGVRGVLDRMQKLDIFEQYNL---EMLKSGDILERDLKKEEARVKKIE 131
DB 1100 LENSQTNEXL--QNELDMLKQNNLKNKEELTKSKELLNLENKKVEELKKEPE 1150

RESULT 12
YB38 MYCPN STANDARD; PRT; 166 AA.
AC P75260;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0134 protein MPN138 (E07_orf166).
GN MPN138 OR MP016.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -!- SIMILARITY: Belongs to the UPF0134 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000003; AAB95664.1; -;

DR PIR; S73342; S73342.
DR InterPro; IPR002862; DUF16.
DR Pfam; PF01519; DUF16; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 166 AA; 19518 MW; BE44F5377B2FA709 CRC64;

Query Match 12.7%; Score 84; DB 1; Length 166;
Best Local Similarity 21.1%; Pred. No. 5.8;
Matches 28; Conservative 30; Mismatches 57; Indels 18; Gaps 3;

QY 17 VSGEDKKHDYQNEF-----DFLLMERIHEQIKKGELALFYLQEQINHFEE 61
DB 31 INPENKSVFESGNGFNQPKLNKPVPVNTVEYVTRKEFNEYKDSNDQRLTKIENKVKLEV 90

QY 62 KPTK--EMKDKIVAEMDTIAMIIDGVRGVLDRMQKLDLIFEQYNLEMLKSGDILER 118
DB 91 KVDKLEEKVDKLEAKVDKLEAKVDKLEAKVDKLEAKVDKLEAKVDKLEAKVDKLEAKV 150

QY 119 DLKKEEARVKKIE 131
DB 151 RLDSIEGRDKIE 163

RESULT 13
UVRC BORBU STANDARD; PRT; 603 AA.
ID UVRC BORBU
AC O51413;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE UVRABC system protein C (UvrC protein) (Excinuclease ABC subunit C).
GN UVRC OR BB0457.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi.";
RL Nature 390:580-586(1997).
CC -!- FUNCTION: The UvrABC repair system catalyzes the recognition and
CC processing of DNA lesions. UvrC both incises the 5' and 3' sides
CC of the lesion. The N-terminal half is responsible for the 3',
CC incision and the C-terminal half is responsible for the 5'
CC incision (By similarity).
CC -!- SUBUNIT: Interacts with uvrB in an incision complex (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the uvrC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001150; AAC66813.1; -;
DR PIR; H70156; H70156.
DR TIGR; BB0457; -;
DR HAMAP; MF 00203; -; 1.
DR InterPro; IPR004791; UvrC.

DR InterPro; IPR001162; UvrC_C.
DR InterPro; IPR000305; UvrC_N.
DR Pfam; PF01541; Exci_endo_N; 1.
DR ProDom; PD005870; UvrC_C; 1.
DR SMART; SM00465; GIYC; 1.
DR TIGRFAMs; TIGR00194; uvrC; 1.
DR PROSITE; PS50164; UVR_C_1; 1.
DR PROSITE; PS50165; UVR_C_2; 1.
KW SOS response; Excision nuclease; DNA repair; DNA recombination;
KW DNA excision; Complete proteome.
SQ SEQUENCE 603 AA; 69832 MW; B9EDDB8B4A63484 CRC64;

Query Match 12.6%; Score 83.5; DB 1; Length 603;
Best Local Similarity 21.1%; Pred. No. 24;
Matches 30; Conservative 28; Mismatches 53; Indels 31; Gaps 5;

QY 1 MKFIIAFVATLAVMTVSGEDKKHDYQNEPFDLLMERIHQIKKGELALFYLOEQ----- 55
Db 291 LQFLIQYTSINMIV-----PDKIHFLKDDIDTKNVEKLINEIKNTKTEIYKETEEILKI 346

QY 56 -----INHFEKPTKEMKD-KIVAEMDTI-----IAMIDVGVGVLDRLMQR 95
Db 347 MEMAISNAELSLREYENKSTKALESKIVLEMDKLPKTIIEGFDIAHLKGQETVASMVTFK 406

QY 96 KDLDFEQYNLEMLKK--SGDI 115
Db 407 MGMPFKENRYRLKLNLSLKGEI 428

RESULT 14
MYSJ_DICDI STANDARD; PRT; 2245 AA.
AC P54697;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Myosin IJ heavy chain.
GN MYOJ.
OS Dictyostelium discoideum (slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX3;
RX MEDLINE=96215148; PubMed=8636147;
RA Hammer J.A. III, Jung G.;
RT "The sequence of the dictyostelium myo J heavy chain gene predicts a
RT novel, dimeric, unconventional myosin with a heavy chain molecular
RT mass of 258 kDa.";
RL J. Biol. Chem. 271:7120-7127(1996).
RN [2]
RP SEQUENCE OF 1-1021 FROM N.A.
RX MEDLINE=97039016; PubMed=8884597;
RA Peterson M.D., Urioste A.S., Titus M.A.;
RT "Dictyostelium discoideum myoJ: a member of a broadly defined myosin
RT V class or a class XI unconventional myosin?";
RL J. Muscle Res. Cell Motil. 17:411-424(1996).
RN [3]
RP SEQUENCE OF 182-298 FROM N.A.
RX MEDLINE=95023928; PubMed=7937787;
RA Titus M.A., Kuspa A., Loomis W.F.;
RT "Discovery of myosin genes by physical mapping in Dictyostelium.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:9446-9450(1994).
CC -!- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN AND HAS ATPASE
CC ACTIVITY THAT IS ACTIVATED BY ACTIN.
CC -!- SUBUNIT: Homodimer.
CC -!- SIMILARITY: Contains 1 myosin-like globular head domain.
CC -!- SIMILARITY: Contains 1 dilute domain.
CC -!- SIMILARITY: Contains 3 IQ domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U42409; AAA85186.1; -;
DR EMBL; L35322; AAA79858.1; -;
DR PIR; T18278; T18278.
DR HSSP; P08799; LMND.
DR DictyDb; DD01095; myoJ.
DR InterPro; IPR002710; DIL.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR001609; myosin_head.
DR InterPro; IPR004009; Myosin_N.
DR Pfam; PF01843; DIL; 1.
DR Pfam; PF00612; IQ; 6.
DR Pfam; PF00063; myosin_head; 2.
DR Pfam; PF02736; Myosin_N; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD003376; DIL; 1.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 3.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 3.
KW Myosin; ATP-binding; Actin-binding; Calmodulin-binding; Repeat;
KW Coiled coil.
FT DOMAIN 1 809 MYOSIN HEAD-LIKE.
FT DOMAIN 824 851 IQ 1.
FT DOMAIN 872 901 IQ 2.
FT DOMAIN 943 972 IQ 3.
FT DOMAIN 973 1812 COILED COIL (POTENTIAL).
FT DOMAIN 2060 2220 DILUTE.
FT NP_BIND 174 181 ATP (POTENTIAL).
FT DOMAIN 669 749 ACTIN-BINDING.
FT DOMAIN 191 191 L -> F (IN REF. 2).
FT CONFLICT 284 284 A -> T (IN REF. 2).
FT CONFLICT 291 291 G -> R (IN REF. 2).
FT CONFLICT 332 347 NKSGCFEIEGVSDDEH -> IEWMFELKVYRMKS (IN
FT REF. 2).
FT CONFLICT 550 550 N -> K (IN REF. 2).
FT CONFLICT 865 866 HH -> QQ (IN REF. 2).
SQ SEQUENCE 2245 AA; 258478 MW; 615E5EF1D1AB45BE CRC64;

Query Match 12.5%; Score 83; DB 1; Length 2245;
Best Local Similarity 27.4%; Pred. No. 1e+02;
Matches 23; Conservative 17; Mismatches 34; Indels 10; Gaps 3;

QY 39 HEQIKKGELALFYLOEQINHFEE--KPTKEMKDKIVAEMDTIIAMIDVGVGVLDRLMQRK 96
Db 1082 HDSIEK-----LQSQFNETEQQLOQFKQSQSEELSKLSKTTQQLDFNKQEFDRLSQER 1134

QY 97 DLD-IFEQYNLEMLKKSGDILERD 119
Db 1135 DTDNTNNQLEIQQLKKANSTLEED 1158

RESULT 15
SYA_THETN STANDARD; PRT; 878 AA.
AC Q8RAH4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alanyl-tRNA synthetase (EC 6.1.1.7) (Alanine--tRNA ligase) (AlaRS).
GN ALAS OR TTE1248.
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336;

RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
CC -!- CATALYTIC ACTIVITY: ATP + L-alanine + tRNA(Ala) = AMP +
CC phosphate + L-alanyl-tRNA(Ala).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE013087; AAM24472.1; -.
DR HAMAP; MF 00036; -. 1.
DR InterPro; IPR003156; DHHA1.
DR InterPro; IPR002318; tRNA-synt_2c.
DR InterPro; IPR006193; tRNA_synt_Ala.
DR Pfam; PF02272; DHHA1; 1.
DR Pfam; PF01411; tRNA-synt_2c; 1.
DR TIGRFAMs; TIGR00344; alaS; 1.
DR PROSITE; PS50860; AA_TRNA_LIGASE_II_ALA; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
SQ SEQUENCE 878 AA; 98363 MW; F2CE98288DC2A244 CRC64;

Query Match 12.5%; Score 82.5; DB 1; Length 878;
Best Local Similarity 29.8%; Pred. No. 42;
Matches 31; Conservative 22; Mismatches 32; Indels 19; Gaps 6;

QY 26 YQNEFDLLMERIHEQIKKGELALF----YLQEQINHFEKPKEMKDKIVAEM-DTII- 79
Db 706 YLNEQKEIL-DRVSETLKAQEKVSVSKIENTLQOSLKD-KEREIEGLKTKIASILAETLID 763

QY 80 --AMIDGVRGVLDRLMQRKDLDFEQYNLEMLKKSGDILERDLK 121
Db 764 SAISVDGVKVIASRV-----EDYDMEALKTLGDILKDLK 798

Search completed: November 17, 2003, 08:23:39
Job time : 18 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2003, 08:21:04 ; Search time 34 Seconds
(without alignments)
1001.851 Million cell updates/sec

Title: US-09-877-160A-1
Perfect score: 662
Sequence: 1 MKFIIAFFVATLAVMTVSGE.....GDILERDLKKEARVKKIEV 132

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	103	15.6	2138	5 Q9XZE3	Q9xze3 amoeba prot
2	94.5	14.3	506	16 Q8E6I3	Q8e6i3 streptococc
3	94.5	14.3	506	16 Q8E128	Q8e128 streptococc
4	94	14.2	860	13 Q42351	Q42351 gallus gall
5	92	13.9	776	16 Q8XKK4	Q8xkk4 clostridium
6	90.5	13.7	920	10 Q8S4W6	Q8s4w6 chlamydomon
7	89	13.4	249	16 Q9ZKG5	Q9zkg5 helicobacte
8	88.5	13.4	558	16 Q8RGM3	Q8rgm3 fusobacteri
9	88.5	13.4	846	4 Q75130	Q75130 homo sapien
10	88.5	13.4	939	16 Q98QI8	Q98qi8 mycoplasma
11	87.5	13.2	459	5 Q04011	Q04011 wuchereria
12	87.5	13.2	623	4 Q9H2G3	Q9h2g3 homo sapien
13	87.5	13.2	2228	5 Q8I1W5	Q8i1w5 plasmodium
14	87	13.1	311	17 Q971V5	Q971v5 sulfolobus
15	86	13.0	249	16 Q25166	Q25166 helicobacte
16	86	13.0	301	16 Q8XL82	Q8xl82 clostridium

17	86	13.0	1158	16 Q8XPE3	Q8xpe3 clostridium
18	86	13.0	3187	11 Q63714	Q63714 rattus norv
19	85	12.8	230	2 Q8GL35	Q8gl35 borrelia bu
20	85	12.8	784	16 Q8ZRB9	Q8zrb9 salmonella
21	85	12.8	784	16 Q8Z8V0	Q8z8v0 salmonella
22	85	12.8	784	16 Q8XE60	Q8xe60 escherichia
23	85	12.8	2413	5 Q8ICH7	Q8ich7 plasmodium
24	84.5	12.8	536	12 Q8JS31	Q8js31 phthorimaea
25	84	12.7	185	16 Q51736	Q51736 borrelia bu
26	84	12.7	1650	5 Q812T7	Q8i2t7 plasmodium
27	83.5	12.6	215	16 Q8EW76	Q8ew76 mycoplasma
28	83.5	12.6	311	6 Q95K18	Q95k18 macaca fasc
29	83	12.5	155	16 Q8EMU6	Q8emu6 oceanobacil
30	83	12.5	559	5 Q02263	Q02263 caenorhabdi
31	83	12.5	951	5 Q96229	Q96229 plasmodium
32	83	12.5	1310	10 Q949K0	Q949k0 lycopersico
33	82.5	12.5	233	2 Q44739	Q44739 borrelia bu
34	82.5	12.5	445	16 Q9PPJ8	Q9ppj8 campylobact
35	82.5	12.5	476	6 Q8WP20	Q8wp20 macaca fasc
36	82.5	12.5	1201	16 Q8DAL0	Q8dal0 vibrio vuln
37	82	12.4	597	5 Q9NCN3	Q9ncn3 plasmodium
38	82	12.4	599	5 Q9NCM5	Q9ncm5 plasmodium
39	82	12.4	599	5 Q9NCM7	Q9ncm7 plasmodium
40	82	12.4	690	5 Q8SSE8	Q8sse8 encephalito
41	82	12.4	1558	5 Q96275	Q96275 plasmodium
42	82	12.4	1716	5 Q8I1M2	Q8i1m2 plasmodium
43	82	12.4	1751	5 Q26194	Q26194 plasmodium
44	82	12.4	1756	5 Q8I1M1	Q8i1m1 plasmodium
45	82	12.4	1756	5 Q8I058	Q8i058 plasmodium

ALIGNMENTS

RESULT 1
Q9XZE3
ID Q9XZE3 PRELIMINARY; PRT; 2138 AA.
AC Q9XZE3;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Myosin heavy chain.
OS Amoeba proteus (Amoeba).
OC Eukaryota; Lobosea; Euamoebida; Amoebidae; Amoeba.
OX NCBI_TaxID=5775;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D;
RX MEDLINE=99082369; PubMed=9864850;
RA Oh S.W., Jeon K.W.;
RT "Characterization of myosin heavy chain and its gene in Amoeba proteus.";
RL J. Eukaryot. Microbiol. 45:600-605(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=D;
RX Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RA EMBL; AF136711; AAD33718.1; -.
DR HSSP; P08799; 1MND.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR001609; myosin_head.
DR InterPro; IPR004009; Myosin_N.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00663; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 1.
SQ SEQUENCE 2138 AA; 244427 MW; DE58B8308639FA71 CRC64;

Query Match 15.6%; Score 103; DB 5; Length 2138;
Best Local Similarity 29.6%; Pred. No. 7.1;
Matches 29; Conservative 24; Mismatches 35; Indels 10; Gaps 3;
QY 35 MERIHEQIKKGELALFYLOEQIN--HFEEKPTKEMKDKIVAEMDTIIAMIDGVRGVLDRMQ 94
Db 1730 LEELEREQLEAEBAALNYL-EEIKKKDLLEINLRKQLDAESE-----ARDKFEQLKN 1780
QY 95 RKDLDFE-QYNLEMLKSGDILERDLKKEEARVKKIE 131
Db 1781 ELERDVADAKHNLEAEKKSRTDAERAKKAEAAQYDELK 1818
RESULT 2
Q8E6I3 PRELIMINARY; PRT; 506 AA.
AC Q8E6I3;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
GN GBS0580.
OS Streptococcus agalactiae (serotype III).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216495;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEM316 / Serotype III;
RX MEDLINE=22242508; PubMed=12354221;
RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Frangeul L.,
RA Msadek T., Zouine M., Couve E., Lalioui L., Poyart C., Trieu-Cuot P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
RT invasive neonatal disease.";
RL Mol. Microbiol. 45:1499-1513 (2002).
DR EMBL; AL766846; CAD46224.1; --
DR Sagalish; gbs0580; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 506 AA; 57621 MW; A617EAC632CF8CDB CRC64;
Query Match 14.3%; Score 94.5; DB 16; Length 506;
Best Local Similarity 26.3%; Pred. No. 7.7;
Matches 25; Conservative 24; Mismatches 33; Indels 13; Gaps 3;
QY 45 GELALFYLOEQIN--HFEEKPTKEMKDKIVAEMDTIIAMIDGVRGVLDRMQKDLDFE 102
Db 240 GELAKYIKKYNINYYIPEENASNKVAKTLADEVGKTAFLSPLEG-----LSKKEMAAGE 294
QY 103 QY-----NLEMLKSGDILERDLKKEEARVKKIE 131
Db 295 DYFSVMRRNLKVLKKTDDVAGKEVAPEEDKTKTVE 329
RESULT 3
Q8E128 PRELIMINARY; PRT; 506 AA.
AC Q8E128;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Zinc ABC transporter, zinc-binding adhesion liprotein.
GN SAG0535.
OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=22222988; PubMed=12200547;
RA Tetelin H., Massignani V., Cieslewicz M.J., Eisen J.A., Peterson S.,
RA Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,

RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,
RA Carly H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni P., Maione D.,
RA Rinaudo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V Streptococcus agalactiae.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396 (2002).
DR EMBL; AE014216; AAM99436.1; --
DR TIGR; SAG0535; --
KW Complete proteome.
SQ SEQUENCE 506 AA; 57621 MW; A617EAC632CF8CDB CRC64;
Query Match 14.3%; Score 94.5; DB 16; Length 506;
Best Local Similarity 26.3%; Pred. No. 7.7;
Matches 25; Conservative 24; Mismatches 33; Indels 13; Gaps 3;
QY 45 GELALFYLOEQIN--HFEEKPTKEMKDKIVAEMDTIIAMIDGVRGVLDRMQKDLDFE 102
Db 240 GELAKYIKKYNINYYIPEENASNKVAKTLADEVGKTAFLSPLEG-----LSKKEMAAGE 294
QY 103 QY-----NLEMLKSGDILERDLKKEEARVKKIE 131
Db 295 DYFSVMRRNLKVLKKTDDVAGKEVAPEEDKTKTVE 329
RESULT 4
Q42351 PRELIMINARY; PRT; 860 AA.
AC Q42351;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Chicken RABAPTIN-5 (NEUROCRESCIN).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Nogawa M., Uyeda A., Nishimune H., Taguchi T.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 79-860 FROM N.A.
RC TISSUE=Muscle, and Telencephalon;
RX MEDLINE=98087215; PubMed=9427343;
RA Nishimune H., Uyeda A., Nogawa M., Fujimori K., Taguchi T.;
RT "Neurorescin: a novel neurite-outgrowth factor secreted by muscle
RT after denervation.";
RL NeuroReport 8:3649-3654 (1997).
DR EMBL; D88828; BAA21785.1; --
DR EMBL; D38038; BAA23805.1; --
DR InterPro; IPR003914; Rabaptin.
DR Pfam; PF03528; Rabaptin; 2.
DR PRINTS; PR01432; RABAPTIN.
SQ SEQUENCE 860 AA; 99414 MW; 202FCF2CA845B2BE CRC64;
Query Match 14.2%; Score 94; DB 13; Length 860;
Best Local Similarity 27.9%; Pred. No. 15;
Matches 41; Conservative 26; Mismatches 46; Indels 34; Gaps 8;
QY 16 TVSGEDKKH-DYQNEFDLFLMERIHEQIKKGELALFYLOEQI-----NHFEKPTKE-- 66
Db 649 SLQGGKSHLVSLQAEDFILPEAAEE---LRELILKYREDITISVRTAADHLEEKLAEL 705
QY 67 -MKDKIVAEM-----DTI-----IAMIDGVRGVLDRMQKDLDFE--Y 104
Db 706 FLKEQIQAEQYLQENIETLQLEIENCKEEIASISLKAELERIKVEKEQLSSQYQLQ 765

QY 105 NLEMLKSGDILERDLKKEEARVKIE 131
Db 766 QLESQETKNTLEBQLKETAAKANLE 792

RESULT 5
Q8XKK4
ID Q8XKK4 PRELIMINARY; PRT; 776 AA.
AC Q8XKK4;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE ATP-dependent protease La.
GN LONA OR CPE1390.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / Type A;
RX PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.,
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater."
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR EMBL; AP003190; BAB81096.1; -.
DR MEROPS; S16.001; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR InterPro; IPR003111; LON.
DR InterPro; IPR001984; Lon_endopep.
DR InterPro; IPR004815; Lon_fam.
DR Pfam; PF00004; AAA; 1.
DR Pfam; PF02190; LON; 1.
DR PRINTS; PR00830; ENDOLAPTASE.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00464; LON; 1.
DR TIGRFAMS; TIGR00763; Lon; 1.
DR PROSITE; PS01046; LON_SER; 1.
KW Protease; Complete proteome.
SQ SEQUENCE 776 AA; 87428 MW; F818519BF7174D19 CRC64;

Query Match 13.9%; Score 92; DB 16; Length 776;
Best Local Similarity 27.7%; Pred. No. 19;
Matches 41; Conservative 21; Mismatches 58; Indels 28; Gaps 7;

QY 1 MKFIIAFFVATLAVMTVSGEDKKHDYQNEFDLLMERIHEQIKKGELALFYLOEQINHFE 60
Db 74 IKQIVKMPKNTIRVLVEGIERAKMD--EFD-----KEELLEASIEKIDIDNEIDHEL 124

QY 61 EKPTKEMKDIIVAEEMDTIIMIDGVGV--LDRLMQRKOL----DIFEQYNLEMLKSGD 114
Db 125 EALSRLKLDFFFEFLD--ITASSGINGVDLFDNLEEKDLNKNVTDLISSVALIKQEDKQD 182

QY 115 ILER-DLK-----KEEARVKIE 131
Db 183 ILQTLDLKRIEKILIFYVKEEIEVAKIE 210

RESULT 6
Q8S4W6
ID Q8S4W6 PRELIMINARY; PRT; 920 AA.
AC Q8S4W6;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Coiled-coil flagellar protein.
GN MBO2.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.

OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11977094;
RA Tam L.W., Lefebvre P.A.;
RT "The Chlamydomonas MBO2 locus encodes a conserved coiled-coil protein
RT important for flagellar waveform conversion."
RL Cell Motil. Cytoskeleton 51:197-212(2002).
DR EMBL; AF394181; AAM15771.1; -.
KW Flagella.
SQ SEQUENCE 920 AA; 104210 MW; 9B24A671A6408830 CRC64;

Query Match 13.7%; Score 90.5; DB 10; Length 920;
Best Local Similarity 23.9%; Pred. No. 31;
Matches 38; Conservative 30; Mismatches 38; Indels 53; Gaps 6;

QY 16 TVSGEDKKHDYQNEFDL-----LMERIHEQ-----IKKGELALFYLOE 54
Db 509 TTEMKDKLVQLQNELDILQNEVGIKDKLQOQHTQHAANFAERDQLRVELGLGMVFRDK 568

QY 55 QINHFEEKPKEMKDIIVAEEMDTIIMIDGVGVLDRLMQRKLDI----- 100
Db 569 Q-----AVVDEQIAEVDKLNAININGCEKEMRLRKQYELVIEARNYTGIMLIDRN 618

QY 101 -----FEQYN-LEMLKSGDILERDLKKEEARVKIE 132
Db 619 DELCVLYEKANILDEVIKSQ-LELMRREDEARLLRLEV 656

RESULT 7
Q9ZKG5
ID Q9ZKG5 PRELIMINARY; PRT; 249 AA.
AC Q9ZKG5;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE Putative PARALOG of HPAA.
GN JHP0971.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori."
RL Nature 397:176-180(1999).
DR EMBL; AB001526; AAD06549.1; -.
KW Complete proteome.
SQ SEQUENCE 249 AA; 28513 MW; EC8D9AFDC99548AA CRC64;

Query Match 13.4%; Score 89; DB 16; Length 249;
Best Local Similarity 28.8%; Pred. No. 10;
Matches 34; Conservative 27; Mismatches 39; Indels 18; Gaps 8;

QY 25 DYQNEFD---FLLMERIHEQIKKGELALFYLOEQINHFEEKPKTKEMDK--IVAEMDTII 79
Db 76 NYQKEFEKSLFLQLSNFE--RKG-----YSVSQFKDVSEIP-QDIKERALLVLRMDGNV 127

QY 80 AMIDGVGVLDRLMQRKLDIFEQYNLEMLK-KSGDILER---DLKKEEARVKIE 132
Db 128 AILEDIVVEESDALSEEKVIDMSSGYLNFVEPKSEDIHSGIDVSKIKAVIERVEL 185

RESULT 8
Q8RGM3
ID Q8RGM3 PRELIMINARY; PRT; 558 AA.

AC Q8RGM3;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE DNA repair protein recN.
GN FN0268.
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11889109;
RA Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fonstein M., Kyrpides N., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586.";
RL J. Bacteriol. 184:2005-2018 (2002).
DR EMBL; AE010538; AAL94474.1; --
DR InterPro; IPR003593; AAA_Atpase.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR004604; RecN.
DR InterPro; IPR002078; Sig54_interact.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMS; TIGR00634; recN; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; 1.
KW Complete proteome.
SQ SEQUENCE 558 AA; 64380 MW; FB47874A8F91B263 CRC64;

Query Match 13.4%; Score 88.5; DB 16; Length 558;
Best Local Similarity 26.6%; Pred. No. 26;
Matches 33; Conservative 26; Mismatches 46; Indels 19; Gaps 5;

QY 13 AVMTVSGEDKKHDYQNEFDLLMERIHEQIKKGELALFYLOEQINHFEE-----KPTKEM 67
Db 247 ALSLTINSIRNIEYLGKYD-----ERYTELAKRIENAYYELEDCAEIEDISKGDVVTEND 302

QY 68 KDKIVAEMDTIIMIDGVRGVLDRLMQ-RKOLDIFEQYNLEMLK--KSGDILERDLKKEE 124
Db 303 LDKIASRMNTLKRKEKYKRTLPELIEYREDLK-----EKLSDDSDGDFKTKELKKEL 355

QY 125 ARVK 128
Db 356 NKIK 359

RESULT 9
O75130
ID O75130 PRELIMINARY; PRT; 846 AA.
AC O75130;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein KIAA0635.
GN KIAA0635.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98403880; PubMed=9734811;
RA Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H.,
RA Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. X.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro."
RL DNA Res. 5:169-176 (1998).

DR EMBL; AB014535; BAA31610.1; --
KW Hypothetical protein.
SQ SEQUENCE 846 AA; 99213 MW; 600B6BEDAA35514F CRC64;

Query Match 13.4%; Score 88.5; DB 4; Length 846;
Best Local Similarity 24.3%; Pred. No. 41;
Matches 27; Conservative 26; Mismatches 43; Indels 15; Gaps 4;

QY 33 LIMERIHE-----QIKKGELALFYLOEQINHFEEKPTKEMKDKIVAEMDTIAMI 82
Db 633 LLEKEIQEHINAHAYESQISSMAKMSRLEELRHQDEKATVLD--LSSLRELICKL 690

QY 83 DGVRGVLDRLMORKDLDFEQ--YNLEMLKKSGDILERDLKKHEARVKKIE 131
Db 691 DSGKDIMTQQLNSKNLE-FERVVVVELENVKSESDLLKKQLSNERHTVKNLE 740

RESULT 10
Q98QI8
ID Q98QI8 PRELIMINARY; PRT; 939 AA.
AC Q98QI8;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Topoisomerase IV subunit A.
GN MYPU_3730.
OS Mycoplasma pulmonis.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2107;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAB CTIP;
RX MEDLINE=21267165; PubMed=11353084;
RA Chambaud I., Heilig R., Ferris S., Barbe V., Samson D., Galisson F.,
RA Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P.C.,
RA Blanchard A.;
RT "The complete genome sequence of the murine respiratory pathogen
RT Mycoplasma pulmonis.";
RL Nucleic Acids Res. 29:2145-2153 (2001).
DR EMBL; AL445564; CAC13546.1; --
DR MypUList; MYPU_3730; --
DR InterPro; IPR002205; DNA_topoisoIV.
DR InterPro; IPR005741; TopoIV_A_Gpos.
DR Pfam; PF00521; DNA_topoisoIV; 1.
DR ProDom; PD000742; DNA_topoisoIV; 1.
DR SMART; SM00434; TOP4c; 1.
DR TIGRFAMS; TIGR01061; parC_gpos; 1.
KW Complete proteome.
SQ SEQUENCE 939 AA; 108050 MW; 4B4830AE99279A8E CRC64;

Query Match 13.4%; Score 88.5; DB 16; Length 939;
Best Local Similarity 24.5%; Pred. No. 45;
Matches 35; Conservative 19; Mismatches 34; Indels 55; Gaps 5;

QY 44 KGELALFYLOEQINHFEEK-----PTKEMKDKIVAEMDTIIA--MIDGVRGVLDLDR--- 91
Db 245 RGRISLTCKYEHIINLPDKNEILLITQIPYGSVKSVKIVRIDEIIVNQKIDGIKEVLDQSDR 304

QY 92 -----LMQRKDLDFEQYN-----LEML 109
Db 305 DGISIAIKLEKKANHESIMKYLFSKTELOVFYNNYNNIAKNGSPTLNLNELLDIYIEHL 364

QY 110 KK-SGDILERDLKKHEARVKKIE 131
Db 365 KTNVTKILEFDLNLKARLLEVE 387

RESULT 11
Q04011
ID Q04011 PRELIMINARY; PRT; 459 AA.
AC Q04011;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)


```

Query Match          13.0%; Score 86; DB 16; Length 249;
Best Local Similarity 29.1%; Pred. No. 18;
Matches 34; Conservative 26; Mismatches 39; Indels 18; Gaps 8;

Qy      26 YONEFD---FLLMERIHEIQIKGELALFYLQEINHFEEKPTKEMDK--IVAEMDTIIA 80
        ||||| : | | : | | | | | : | | : | | : | | : | | : | |
Db      77 YQKEFEKSIFQLSSFLE--RKG-----YSVSQFKDASEIP-QDIKEKALLVLRMDGVA 128

Qy      81 MIDGVRGVLDRLMQRKDLDFEQY-NLEMLK-KSGDILER---DLKKEEARVKKIEV 132
        :: : | | : | : | | : | | : | | : | | : | | : | | : | |
Db     129 ILEDIVERSDALSEEKVDMSSGYLNLFVEPKSEDIITHSFGIDVSKIKAVIERVEL 185

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2003, 08:17:29 ; Search time 41 Seconds
(without alignments)
511.022 Million cell updates/sec

Title: US-09-877-160A-1
Perfect score: 662
Sequence: 1 MKFIIAFFVATLAVMTVSGE.....GDILDLKKEARVKKIEV 132

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 19Jun03:*

1:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
24:	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	657	99.2	132	20	AA25584 D. pteronyssinus a
2	273.5	41.3	134	22	AAE02255 Domestic mite Bt5
3	273.5	41.3	134	23	AAE19279 Blomia tropicalis
4	273.5	41.3	134	23	AAE19386 Blomia tropicalis
5	268.5	40.6	134	22	AAE02256 Domestic mite Bt5
6	268.5	40.6	134	23	AAE19280 Blomia tropicalis
7	262	39.6	127	22	AAE02257 Domestic mite Bt5
8	262	39.6	127	23	AAE19278 Blomia tropicalis
9	261	39.4	117	22	AAE02251 Colombian mite Bt5

10	261	39.4	117	23	AAE19273 Blomia tropicalis
11	261	39.4	117	23	AAE19275 Blomia tropicalis
12	259	39.1	117	22	AAE02248 Domestic mite Bt5
13	259	39.1	117	22	AAE02249 Domestic mite Bt5
14	259	39.1	117	22	AAE02253 Domestic Bt5 polym
15	259	39.1	117	23	AAE19271 Blomia tropicalis
16	259	39.1	117	23	AAE19276 Blomia tropicalis
17	253	38.2	117	22	AAE02250 Colombian mite Bt5
18	253	38.2	117	22	AAE02252 Singapore mite Bt5
19	253	38.2	117	23	AAE19272 Blomia tropicalis
20	253	38.2	117	23	AAE19274 Blomia tropicalis
21	176	26.6	84	22	AAE02258 Domestic mite Bt5
22	176	26.6	84	23	AAE19281 Streptococcus poly
23	94.5	14.3	506	23	ABP28185 Streptococcus poly
24	94.5	14.3	506	23	ABP29862 Chicken neurite pr
25	94	14.2	782	23	AAU98072 Chicken derived ne
26	91	13.7	687	17	AAU97123 H. pylori ORF 01ae
27	89	13.4	249	18	AAW55445 H. pylori outer me
28	89	13.4	249	20	AAU17212 H. pylori ORF hp3e
29	89	13.4	261	18	AAW55332 Inner centromere p
30	88	13.3	410	23	ABP43547 Intracellular traf
31	87.5	13.2	584	22	ABU53081 Human tumour marke
32	87.5	13.2	623	23	ABB77434 Intracellular traf
33	87.5	13.2	711	22	ABU53080 Transport and traf
34	87.5	13.2	766	22	ABU53068 Intracellular traf
35	87.5	13.2	907	22	ABU53073 Intracellular traf
36	87.5	13.2	941	22	ABU53072 Intracellular traf
37	87.5	13.2	953	22	ABU53069 Intracellular traf
38	87.5	13.2	2230	24	ABU07445 Protein differenti
39	87	13.1	298	23	AAE25655 Human PIBF 35kDa p
40	87	13.1	308	23	AAE25662 Human PIBF 37kDa p
41	87	13.1	612	22	AAW79138 Human protein SEQ
42	87	13.1	690	22	AAW80122 Human protein SEQ
43	87	13.1	757	23	AAE25653 Human progesterone
44	87	13.1	757	23	AAE25671 Human progesterone
45	86	13.0	132	23	ABU51422 Helicobacter pylor

ALIGNMENTS

RESULT 1
AA25584
ID AA25584 standard; protein; 132 AA.
XX
AC AA25584;
XX
DT 30-SEP-1999 (first entry)
XX
DE D. pteronyssinus allergen Der p 5 protein fragment.
XX
KW Major histocompatibility complex; class II; desensitising; human;
KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
KW chiromidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; guinea pig;
KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity.
XX
OS Dermatophagoides pteronyssinus.
XX
PN WO9934826-A1.
XX
PD 15-JUL-1999.
XX
PF 11-JAN-1999; 99WO-GB00080.
XX
PR 21-SEP-1998; 98GB-0020474.
XX
PR 09-JAN-1998; 98GB-0000445.
XX
PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
PI Kay AB, Larche M;
XX

DR WPI; 1999-458255/38.
XX Desensitizing patients to polypeptide allergens
XX
PS Example 6; Page 51; 117pp; English.
XX
CC This invention describes a novel method of desensitizing a patient to a
CC polypeptide allergen and comprises administering to the patient a peptide
CC derived from the allergen where restriction to a MHC Class II molecule
CC possessed by the patient can be demonstrated for the peptide and the
CC peptide is able to induce a late phase response in an individual who
CC possesses the MHC Class II molecule. The methods can be used for
CC desensitizing patients to allergens present in e.g. grass, tree and weed
CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
CC Teribrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
CC produce immunological vaccines which may be used to prevent and/or treat
CC conditions involving hypersensitivity to allergens. This sequence
CC represents the house dust mite (Dermatophagoides pteronyssinus) allergen
CC Der p 5.
XX
SQ Sequence 132 AA;

Query Match 99.2%; Score 657; DB 20; Length 132;
Best Local Similarity 99.2%; Pred. No. 7.7e-59;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MKFTIAFFVATLAVMTVSGEDKKHDYQNEFFDLMERHEIQKKGELALFYLOEQINHFE 60
Db 1 MKFTIAFFVATLAVMTVSGEDKKHDYQNEFFDLMERHEIQKKGELALFYLOEQINHFE 60
QY 61 EKPTKEMKDKIVAEMDTIAMIADGVGVLDRLMQRKDLDFEQYNLEMLKSGDILERDL 120
Db 61 EKPTKEMKDKIVAEMDTIAMIADGVGVLDRLMQRKDLDFEQYNLEMLKSGDILERDL 120
QY 121 KKEEARVKKIEV 132
Db 121 KKEEARVKKIEV 132

RESULT 2
AAE02255
ID AAE02255 standard; Protein; 134 AA.
AC AAE02255;
XX
DT 31-JUL-2001 (first entry)
XX
DE Domestic mite Bt5 polymorphic amino acid variant no.7.
XX
KW Mite; immunogenic protein; Bt allergen; therapy; atopic dermatitis;
KW immediate hypersensitivity; systemic anaphylaxis; allergic rhinitis;
KW asthma; anti-allergic; anti-inflammatory; immunosuppressive.
XX
OS Blomia tropicalis.
XX
PN WO200130817-A1.
XX
PD 03-MAY-2001.
XX
PF 10-OCT-2000; 2000WO-AU01227.
XX
PR 26-OCT-1999; 99SG-0005313.
PR 18-JUL-2000; 2000AU-0008842.
PR 18-JUL-2000; 2000AU-0008844.
PR 18-JUL-2000; 2000AU-0008845.
XX
PA (UYSI-) UNIV SINGAPORE NAT.
XX
PI Chua KY, Cheong N, Lee BW;

XX WPI; 2001-308609/32.
DR N-PSDB; AAD06260.
XX
PT Novel immunogenic protein derived from house mite, Blomia tropicalis
PT useful for treating and diagnosing conditions involving induction of
PT immuneresponse to mite, such as allergic asthma, atopic dermatitis,
PT rhinitis -
XX
PS Example 46; Page 219-220; 230pp; English.
XX
CC The present invention relates to immunogenic proteins, referred as Bt
CC allergen, is derived from domestic mite, Blomia tropicalis. The specific
CC Bt allergens of the invention includes Bt11, Bt10, Bt5 and BtA2. The
CC immunogenic protein is useful for preventing, reducing or ameliorating
CC Blomia tropicalis hypersensitivity condition, such as atopic dermatitis,
CC immediate hypersensitivity, systemic anaphylaxis, allergic rhinitis or
CC asthma and for modulating an immune response directed to Bt allergen in
CC a subject. The Bt allergens are also useful for detecting antibody
CC directed to all or a part of Bt allergen in a biological sample from a
CC subject. Antibodies to Bt allergens are also used as therapeutic or
CC diagnostic agents, to screen Bt immunoassays and as antagonists to
CC inhibit Bt activity under circumstances where temporary hypersensitivity
CC inhibition is required. The present sequence is Bt5 allergen
CC polymorphic amino acid variant.
XX
SQ Sequence 134 AA;

Query Match 41.3%; Score 273.5; DB 22; Length 134;
Best Local Similarity 42.5%; Pred. No. 5.4e-20;
Matches 57; Conservative 32; Mismatches 40; Indels 5; Gaps 2;
QY 1 MKF---IIAFFVATLAVMTVSGEDKKHDYQNEFFDLMERHEIQKKGELALFYLOEQIN 57
Db 1 MKFAVLVIACFAA--SVLAQEHKPKDDFRNEFDHLLIEQANHAIEKGEHQLLYLQHOLD 58
QY 58 HFEEKPTKEMKDKIVAEMDTIAMIADGVGVLDRLMQRKDLDFEQYNLEMLKSGDILE 117
Db 59 ELNENKSKELQEKIILELDVVCAMIEGAQGALERELKRTLNLILFRFYEEAQTLSKILL 118
QY 118 RDLKKEEARVKKIE 131
Db 119 KDLKETEQVKVDIQ 132

RESULT 3
AAE19279
ID AAE19279 standard; Protein; 134 AA.
XX
AC AAE19279;
XX
DT 21-MAY-2002 (first entry)
XX
DE Blomia tropicalis Bt5 polymorphic amino acid variant no. (2)7.
XX
KW Immunogenic polymorphic Bt5 protein; domestic mite allergen; rhinitis;
KW hypersensitivity; immune response; allergy; asthma; atopic dermatitis;
KW pharmaceutical composition; anti-allergic; antiasthmatic; dermatological;
KW anti-inflammatory; vaccine; gene therapy; mutant; mutein; variant.
XX
OS Blomia tropicalis.
OS Synthetic.
XX
PN WO200206323-A1.
XX
PD 24-JAN-2002.
XX
PF 10-OCT-2000; 2000WO-AU01228.
XX
PR 18-JUL-2000; 2000AU-0008845.
XX
PA (UYSI-) UNIV SINGAPORE NAT.

XX Blomia tropicalis.
OS
XX
FH Key Location/Qualifiers
FT Misc-difference 17 /note= "Encoded by TGG"
FT Misc-difference 34 /note= "Encoded by ATT"
FT Misc-difference 35 /note= "Encoded by CAC"
FT
XX WO200130817-A1.
XX
XX PD 03-MAY-2001.
XX
XX PF 10-OCT-2000; 2000WO-AU01227.
XX
XX PR 26-OCT-1999; 99SG-0005313.
XX PR 18-JUL-2000; 2000AU-0008842.
XX PR 18-JUL-2000; 2000AU-0008844.
XX PR 18-JUL-2000; 2000AU-0008845.
XX
XX PA (UYSI-) UNIV SINGAPORE NAT.
XX
XX PI Chua KY, Cheong N, Lee BW;
XX
XX WPI; 2001-308609/32.
XX
XX DR N-PSDB; AAD06248, AAD06249, AAD06250, AAD06254, AAD06256, AAD06257.
XX
XX PT Novel immunogenic protein derived from house mite, Blomia tropicalis
PT useful for treating and diagnosing conditions involving induction of
PT immuneresponse to mite, such as allergic asthma, atopic dermatitis,
PT rhinitis -
XX
XX Example 46; Page 209; 230pp; English.
XX
XX CC The present invention relates to immunogenic proteins, referred as Bt
CC allergen, is derived from domestic mite, Blomia tropicalis. The specific
CC Bt allergens of the invention includes Bt11, Bt10, Bt5 and BtA2. The
CC immunogenic protein is useful for preventing, reducing or ameliorating
CC Blomia tropicalis hypersensitivity condition such as atopic dermatitis,
CC immediate hypersensitivity, systemic anaphylaxis, allergic rhinitis or
CC asthma and for modulating an immune response directed to Bt allergen in
CC a subject. The Bt allergens are also useful for detecting antibody
CC directed to all or a part of Bt allergen in a biological sample from a
CC subject. Antibodies to Bt allergens are also used as therapeutic or
CC diagnostic agents, to screen Bt immunoassays and as antagonists to
CC inhibit Bt activity under circumstances where temporary hypersensitivity
CC inhibition is required. The present sequence is mite Bt5 polymorphic
CC protein.
XX
XX SQ Sequence 117 AA;
Query Match 39.1%; Score 259; DB 22; Length 117;
Best Local Similarity 44.5%; Pred. NO. 1.4e-18;
Matches 49; Conservative 28; Mismatches 33; Indels 0; Gaps 0;
QY 22 KKHQYQNEFDLLMERIHEQIKKGELALFYLOEQINHFEEKPTKEMKDKIVAEMDTIIAM 81
Db 6 KKDDFRNEFDLLIEQANHAIEKGEHQLLYLQHQDLNENKSKELQEKIIRLDVVCAM 65
QY 82 IDGVRGVLDRLMQRKDLDFEQYNLEMLKKSGDILERDLKKEARVKKIE 131
Db 66 IEGAQGALERELKRTDLNILERFNFEYEAQTLSKILLKDLKETEOKVKDIO 115
RESULT 14
AAE02253
ID AAE02253 standard; Protein; 117 AA.
XX
XX AC AAE02253;
XX
DT 31-JUL-2001, (first entry)

XX Domestic Bt5 polymorphic amino acid sequence no: 11.
DE
XX
XX KW Mite; immunogenic protein; Bt allergen; therapy; atopic dermatitis;
KW immediate hypersensitivity; systemic anaphylaxis; allergic rhinitis;
KW asthma; antiallergic; antiinflammatory; immunosuppressive.
XX
XX OS Blomia tropicalis.
XX
XX PN WO200130817-A1.
XX
XX PD 03-MAY-2001.
XX
XX PF 10-OCT-2000; 2000WO-AU01227.
XX
XX PR 26-OCT-1999; 99SG-0005313.
XX PR 18-JUL-2000; 2000AU-0008842.
XX PR 18-JUL-2000; 2000AU-0008844.
XX PR 18-JUL-2000; 2000AU-0008845.
XX
XX PA (UYSI-) UNIV SINGAPORE NAT.
XX
XX PI Chua KY, Cheong N, Lee BW;
XX
XX WPI; 2001-308609/32.
XX
XX DR N-PSDB; AAD06258.
XX
XX PT Novel immunogenic protein derived from house mite, Blomia tropicalis
PT useful for treating and diagnosing conditions involving induction of
PT immuneresponse to mite, such as allergic asthma, atopic dermatitis,
PT rhinitis -
XX
XX PS Example 46; Page 212; 230pp; English.
XX
XX CC The present invention relates to immunogenic proteins, referred as Bt
CC allergen, is derived from domestic mite, Blomia tropicalis. The specific
CC Bt allergens of the invention includes Bt11, Bt10, Bt5 and BtA2. The
CC immunogenic protein is useful for preventing, reducing or ameliorating
CC Blomia tropicalis hypersensitivity condition such as atopic dermatitis,
CC immediate hypersensitivity, systemic anaphylaxis, allergic rhinitis or
CC asthma and for modulating an immune response directed to Bt allergen in
CC a subject. The Bt allergens are also useful for detecting antibody
CC directed to all or a part of Bt allergen in a biological sample from a
CC subject. Antibodies to Bt allergens are also used as therapeutic or
CC diagnostic agents, to screen Bt immunoassays and as antagonists to
CC inhibit Bt activity under circumstances where temporary hypersensitivity
CC inhibition is required. The present sequence is mite Bt5 polymorphic
CC protein.
XX
XX SQ Sequence 117 AA;
Query Match 39.1%; Score 259; DB 22; Length 117;
Best Local Similarity 44.5%; Pred. NO. 1.4e-18;
Matches 49; Conservative 28; Mismatches 33; Indels 0; Gaps 0;
QY 22 KKHQYQNEFDLLMERIHEQIKKGELALFYLOEQINHFEEKPTKEMKDKIVAEMDTIIAM 81
Db 6 KKDDFRNEFDLLIEQANHAIEKGEHQLLYLQHQDLNENKSKELQEKIIRLDVVCAM 65
QY 82 IDGVRGVLDRLMQRKDLDFEQYNLEMLKKSGDILERDLKKEARVKKIE 131
Db 66 IEGAQGALERELKRTDLNILERFNFEYEAQTLSKILLKDLKETEOKVKDIO 115
RESULT 15
AAE19271
ID AAE19271 standard; Protein; 117 AA.
XX
XX AC AAE19271;
XX
DT 21-MAY-2002 (first entry)
XX
XX DE Blomia tropicalis Bt5 amino acid sequence (ref 12).

XX Immunogenic polypeptide Bt5 protein; domestic mite allergen; rhinitis;
KW hypersensitivity; immune response; allergy; asthma; atopic dermatitis;
KW pharmaceutical composition; antiallergic; antiasthmatic; dermatological;
KW antiinflammatory; vaccine; gene therapy.
XX Blomia tropicalis.
OS WO200206323-A1.
XX PN 24-JAN-2002.
XX PD 10-OCT-2000; 2000WO-AU01228.
XX PF 18-JUL-2000; 2000AU-0008845.
XX PR (UYSI-) UNIV SINGAPORE NAT.
XX PA Chua KY, Cheong N, Lee BW, Liew LN;
XX PI WPI; 2002-195803/25.
XX DR N-PSDB; AAD30694.
XX PT Novel immunogenic protein isolated from domestic mite allergen, Blomia
PT tropicalis, useful for preventing, reducing and ameliorating a
PT polypeptide Bt5 hypersensitivity -
XX PS Disclosure; Page 77-78; 101pp; English.
XX CC The present invention relates to an isolated immunogenic polypeptide Bt5
CC protein derived from domestic mite allergen, Blomia tropicalis (Bt). Bt5
CC is useful for preventing, reducing and ameliorating a polypeptide Bt5
CC hypersensitivity. Bt5 is used as vaccine and for modulating an immune
CC response. Bt5 is useful for detecting antibody directed to all or part of
CC polypeptide Bt5 in a biological sample, from a subject, by contacting the
CC biological sample with Bt5, its derivative, homologue, analogue, mimetic
CC or chemical equivalent, to form an antibody-protein complex and detecting
CC the complex. Bt5 is useful for the prophylactic treatment of an allergic
CC condition (including allergic asthma, atopic dermatitis and/or rhinitis)
CC in an individual, e.g. human or animal, by administering Bt5 (preferably
CC in the form of an eukaryotic expression plasmid vector administered in a
CC pharmaceutical composition comprising a normal saline or liposome as
CC carrier), its derivative or homolog, where the airway hyper-reactivity or
CC airway inflammation is prevented. Bt5 gene is useful in gene therapy. The
CC present sequence is Blomia tropicalis Bt5 amino acid sequence (ref 12).
XX SQ Sequence 117 AA;
Query Match 39.1%; Score 259; DB 23; Length 117;
Best Local Similarity 44.5%; Pred. No. 1.4e-18;
Matches 49; Conservative 28; Mismatches 33; Indels 0; Gaps 0;
QY 22 KKHDYQNEFDLLMERIHEQIKKGELALFYLQEQINHFEEKPTKEMKDKIVAEMDTIAM 81
Db 6 KKDDFNEFDHLLIEQANHAIEKGEHQLLYLQHQDLNENKSKELQEKIIRLDVVYCAM 65
QY 82 IDGVRGVLDRLMQKDLDFEQYNLEMLKKSGLDILERDLKKEARVKIE 131
Db 66 IEQAQALERELKRTDLNILERFNYEAAQTLSKILLKDLKETEKVQKDIQ 115

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2003, 08:24:35 ; Search time 29 Seconds
(without alignments)
830.960 Million cell updates/sec

Title: US-09-877-160A-1
Perfect score: 662
Sequence: 1 MKFIIAFFVATLAVMTVSGE.....GDILERDLKKEEARVKKIEV 132

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9:	/cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10:	/cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13:	/cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16:	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	657	99.2	132	10	US-09-877-160-1
2	657	99.2	132	11	US-09-847-208-83
3	86	13.0	249	10	US-09-881-752A-264
4	86	13.0	249	12	US-10-080-113-1
5	82	12.4	379	12	US-10-100-699-2
6	82	12.4	380	9	US-09-134-333-13
7	82	12.4	1786	10	US-09-742-096-3
8	80.5	12.2	676	12	US-10-256-250-16
9	80.5	12.2	709	12	US-10-256-250-15
10	80.5	12.2	725	11	US-09-978-309A-47
11	80.5	12.2	725	12	US-10-256-250-14
12	80.5	12.2	2383	15	US-10-082-830-260
13	80	12.1	1938	15	US-10-171-311-164
14	80	12.1	1945	11	US-09-927-597-2
15	80	12.1	1972	15	US-10-171-311-162

16	80	12.1	1979	11	US-09-927-597-4	Sequence 4, Appli
17	78	11.8	862	15	US-10-177-293-391	Sequence 391, App
18	77.5	11.7	475	10	US-09-738-973-193	Sequence 193, App
19	77.5	11.7	475	10	US-09-854-133-193	Sequence 193, App
20	77.5	11.7	475	15	US-10-144-649A-193	Sequence 193, App
21	77.5	11.7	1055	9	US-09-815-242-11217	Sequence 11217, A
22	77	11.6	380	9	US-09-134-333-12	Sequence 12, Appli
23	76.5	11.6	107	15	US-10-023-066A-92	Sequence 92, Appli
24	76.5	11.6	419	11	US-09-863-049A-2	Sequence 2, Appli
25	76.5	11.6	419	12	US-09-851-673-2	Sequence 3, Appli
26	76	11.5	431	12	US-10-254-995-3	Sequence 3, Appli
27	75.5	11.4	234	11	US-09-819-142-16	Sequence 16, Appli
28	75.5	11.4	234	11	US-09-934-455-358	Sequence 358, App
29	75.5	11.4	967	15	US-10-197-666A-90	Sequence 90, Appli
30	75.5	11.4	1591	15	US-10-197-666A-92	Sequence 92, Appli
31	75	11.3	1707	15	US-10-128-714-3556	Sequence 3556, Ap
32	75	11.3	1750	15	US-10-128-714-8556	Sequence 8556, Ap
33	74.5	11.3	488	12	US-10-032-585-7710	Sequence 7710, Ap
34	74.5	11.3	821	10	US-09-884-001-18	Sequence 18, Appli
35	74.5	11.3	975	15	US-10-097-340-31	Sequence 31, Appli
36	74	11.2	281	12	US-10-080-608A-63	Sequence 63, Appli
37	74	11.2	307	11	US-09-298-523B-7	Sequence 7, Appli
38	74	11.2	715	10	US-09-870-759-86	Sequence 86, Appli
39	74	11.2	715	12	US-09-751-708A-86	Sequence 86, Appli
40	74	11.2	2125	10	US-09-919-172-29	Sequence 29, Appli
41	74	11.2	2649	12	US-10-205-219-169	Sequence 169, App
42	73.5	11.1	208	10	US-09-925-637-52	Sequence 52, Appli
43	73.5	11.1	208	15	US-10-084-205-52	Sequence 52, Appli
44	73.5	11.1	251	12	US-10-254-995-4	Sequence 4, Appli
45	73.5	11.1	254	11	US-09-056-019-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-09-877-160-1
; Sequence 1, Application US/09877160
; Publication No. US20020197268A1
; GENERAL INFORMATION:
; APPLICANT: Ching-Hsaing, Hsu
; APPLICANT: Cheng, Winston T. K.
; TITLE OF INVENTION: ALLERGEN-CONTAINING MILK FOR ALLERGY
; TITLE OF INVENTION: TREATMENT
; FILE REFERENCE: 12774-003001
; CURRENT APPLICATION NUMBER: US/09/877,160
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus
US-09-877-160-1

Query Match	99.2%;	Score 657;	DB 10;	Length 132;
Best Local Similarity	99.2%;	Pred. No. 2.9e-59;		
Matches 131;	Conservative	0;	Mismatches 1;	Indels 0; Gaps 0;
QY	1	MKFIIAFFVATLAVMTVSGEDKKHDYQNEFDLLMERIHEQIKKGELALFYLQEQINHFE	60	
Db	1	MKFIIAFFVATLAVMTVSGEDKKHDYQNEFDLLMERIHEQIKKGELALFYLQEQINHFE	60	
QY	61	EKPTKEMKDKIVAEMDTIIAMIDGVRGVLDRLMQRKLDIFEQYNLEMLKKSGLERDL	120	
Db	61	EKPTKEMKDKIVAEMDTIIAMIDGVRGVLDRLMQRKLDIFEQYNLEMLKKSGLERDL	120	
QY	121	KKEEARVKKIEV 132		
Db	121	KKEEARVKKIEV 132		

RESULT 2

```
US-09-847-208-83
; Sequence 83, Application US/09847208
; Publication No. US20030082190A1
; GENERAL INFORMATION:
; APPLICANT: Saxon, Andrew
; APPLICANT: Zhang, Ke
; APPLICANT: Zhu, Daocheng
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
; TITLE OF INVENTION: IGE-MEDIATED ALLERGIC DISEASES
; FILE REFERENCE: UC67.002A
; CURRENT APPLICATION NUMBER: US/09/847,208
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 83
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Dermatophagoides pteronyssinus (House-dust mite)
US-09-847-208-83

Query Match      99.2%; Score 657; DB 11; Length 132;
Best Local Similarity 99.2%; Pred. No. 2.9e-59;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKFIIAFFVATLAVMTVSGEDKKHDYQNEBDFLLMERIHEQIKKGELALFYLOEQINHF 60
Db 1 MKFIIAFFVATLAVMTVSGEDKKHDYQNEBDFLLMERIHEQIKKGELALFYLOEQINHF 60

QY 61 EKPTKEMKDKIVAEMDTIIAMIDGVRGVLDRMQKDLDFEQYNLEMLKKSGDILRDL 120
Db 61 EKPTKEMKDKIVAEMDTIIAMIDGVRGVLDRMQKDLDFEQYNLEMLKKSGDILRDL 120

QY 121 KKEEARVKKIEV 132
Db 121 KKEEARVKKIEV 132

RESULT 3
US-09-881-752A-264
; Sequence 264, Application US/09881752A
; Patent No. US20020115078A1
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20020115078A1 Helicobacter Polypeptides in the
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 06132/041002
; CURRENT APPLICATION NUMBER: US/09/881,752A
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 08/833,457
; PRIOR FILING DATE: 1997-04-01
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 264
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-881-752A-264

Query Match      13.0%; Score 86; DB 10; Length 249;
Best Local Similarity 29.1%; Pred. No. 0.68;
Matches 34; Conservative 26; Mismatches 39; Indels 18; Gaps 8;

QY 26 YQNEFD---FLLMERIHEQIKKGELALFYLOEQINHFEEKPTKEMKDK--IVAEMDTIIA 80
Db 77 YQKEFEKSLFLQLSSFLE--RKG-----YVSQPKDASEIP-QDIKEKALLVLRMDGNA 128

QY 81 MIDGVRGVLDRMQKDLDFEQY-NLEMLK-KSGDILER---DLKKEEARVKKIEV 132
Db 81 MIDGVRGVLDRMQKDLDFEQY-NLEMLK-KSGDILER---DLKKEEARVKKIEV 132
```

```
Db 129 ILEDIVEESDALSEEKVIDMSSGYLNLNFVEPKSEDIHSFGIDVSKIKAVIERVEL 185

RESULT 4
US-10-080-113-1
; Sequence 1, Application US/10080113
; Publication No. US20030166027A1
; GENERAL INFORMATION:
; APPLICANT: SACHS, GEORGE
; APPLICANT: VOLAND, PETRA
; TITLE OF INVENTION: COMPOSITIONS, TEST KITS AND METHODS FOR DETECTING
; TITLE OF INVENTION: HELICOBACTER PYLORI
; FILE REFERENCE: 626 06 PA
; CURRENT APPLICATION NUMBER: US/10/080,113
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-10-080-113-1

Query Match      13.0%; Score 86; DB 12; Length 249;
Best Local Similarity 29.1%; Pred. No. 0.68;
Matches 34; Conservative 26; Mismatches 39; Indels 18; Gaps 8;

QY 26 YQNEFD---FLLMERIHEQIKKGELALFYLOEQINHFEEKPTKEMKDK--IVAEMDTIIA 80
Db 77 YQKEFEKSLFLQLSSFLE--RKG-----YVSQPKDASEIP-QDIKEKALLVLRMDGNA 128

QY 81 MIDGVRGVLDRMQKDLDFEQY-NLEMLK-KSGDILER---DLKKEEARVKKIEV 132
Db 129 ILEDIVEESDALSEEKVIDMSSGYLNLNFVEPKSEDIHSFGIDVSKIKAVIERVEL 185

RESULT 5
US-10-100-699-2
; Sequence 2, Application US/10100699
; Publication No. US20030157650A1
; GENERAL INFORMATION:
; APPLICANT: Lanar, David E.
; APPLICANT: Dutta, Sheetij
; APPLICANT: Ware, Lisa A.
; TITLE OF INVENTION: Recombinant P. vivax Merozoite Protein-1 p42 Vaccine
; FILE REFERENCE: 003/248/SAP
; CURRENT APPLICATION NUMBER: US/10/100,699
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: US 60/277,002
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 2
; LENGTH: 379
; TYPE: PRT
; ORGANISM: P. vivax
US-10-100-699-2

Query Match      12.4%; Score 82; DB 12; Length 379;
Best Local Similarity 25.0%; Pred. No. 2.9;
Matches 40; Conservative 26; Mismatches 56; Indels 38; Gaps 6;

QY 8 PVATLAVMTVSGEDKKHDYQNEBDFLLMERIHEQIKKGELALFYLOEQINHF 60
Db 28 YLKPLAGMYTIKKQLENHVNAFNTNITMDLSRLKKRNYFLEVLNSDLNPFKYSSGEY 87

QY 61 --EKPTK----EMKDKIVAEMDTIIAMI-----DGVR-----GVLDR 92
Db 88 IKDPYKLLDLLEKKKKLIGSYKIGASIDMDLATANDGVTYNKMGEYKTHLDGVKTEI 147

QY 93 MQRKDLDFEQYNLEMLKKSGDILRDLKKEEARVKKIEV 132
Db 148 KKVED-DIKQD--EELKGLGNVNSQDSKNEFIKKAEL 184
```

```

; SEQ ID NO 3
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: P. falciparum
US-09-742-096-3

Query Match      12.4%; Score 82; DB 10; Length 1786;
Best Local Similarity 24.8%; Pred. No. 21;
Matches 27; Conservative 29; Mismatches 41; Indels 12; Gaps 4;

QY 28 NEEDFLMERIHEQIKKGELALFYLOEQINHFEEKPTKEMKDKIVAEMDTIIAMID---G 84
Db 1237 NDITSKLIETQF-LNEVEADLIKMEKLEKALSSEDSKEIIDAKDDTLEKVIIEHD 1295

QY 85 VRGVLDRLMQRKDLDFEQYNLEMLKKSGDI--LERDLKKEEARVKKIE 131
Db 1296 ITTLDDEVVELKDVE-----EDKIEKVSDLDLEEDILKEVKEIKELE 1338

RESULT 8
US-10-256-250-16
; Sequence 16, Application US/10256250
; Publication No. US20030170755A1
; GENERAL INFORMATION:
; APPLICANT: Schmitt, Michael
; TITLE OF INVENTION: Tumor-Associated Antigen RHAMM
; FILE REFERENCE: L0461/70143 (JRV/NXA)
; CURRENT APPLICATION NUMBER: US/10/256,250
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/324,989
; PRIOR FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 676
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-256-250-16

Query Match      12.2%; Score 80.5; DB 12; Length 676;
Best Local Similarity 23.6%; Pred. No. 8.6;
Matches 35; Conservative 26; Mismatches 34; Indels 53; Gaps 6;

QY 20 EDKKHQYQ-----EFDFLMEIRIHEQIKKGELALFYLOEQ----- 55
Db 298 ETEKEDHVNREHNENLNAEQNLQKFILEQREHEKLOQKELQIDSLLOQKEKLSSSL 357

QY 56 ---INHFEKEPTKE---MKDKIVAEMDTIIAMIDGVRGVLDRLMQR-----KDLDF 101
Db 358 HQKLCSPQEQEEMVKEKNLFEEELKQTLDE-----LDKLOQKEEQAEERLVKQLEEE 406

QY 102 EQYNLEMLKKSGDILERDLKKEEARVKK 129
Db 407 AKSRAEELK---LLEKLGKGAELK 430

RESULT 9
US-10-256-250-15
; Sequence 15, Application US/10256250
; Publication No. US20030170755A1
; GENERAL INFORMATION:
; APPLICANT: Schmitt, Michael
; TITLE OF INVENTION: Tumor-Associated Antigen RHAMM
; FILE REFERENCE: L0461/70143 (JRV/NXA)
; CURRENT APPLICATION NUMBER: US/10/256,250
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/324,989
; PRIOR FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 709
; TYPE: PRT

US-09-134-333-13
; Sequence 13, Application US/09134333
; Patent No. US20020076403A1
; GENERAL INFORMATION:
; APPLICANT: LONGACRE-ANDRE, SHIRLEY
; APPLICANT: ROTH, CHARLES
; APPLICANT: NATO, FARIDABANO
; APPLICANT: BARNWELL, JOHN
; APPLICANT: MENDIS, KAMINI
; TITLE OF INVENTION: RECOMBINANT PROTEIN CONTAINING A C-TERMINAL FRAGMENT OF
; FILE REFERENCE: 0660-0135-0XCIP
; CURRENT APPLICATION NUMBER: US/09/134,333
; CURRENT FILING DATE: 1999-04-18
; EARLIER APPLICATION NUMBER: PCT/FR97/00290
; EARLIER FILING DATE: 1997-02-14
; EARLIER APPLICATION NUMBER: FR96/01822
; EARLIER FILING DATE: 1996-02-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Plasmodium vivax-like sp.
; FEATURE:
; OTHER INFORMATION: Amino Acids 1-140-REGION I
; FEATURE:
; OTHER INFORMATION: Amino Acids 141-178-REGION II
; FEATURE:
; OTHER INFORMATION: Amino Acids 179-283-REGION III
; FEATURE:
; OTHER INFORMATION: Amino Acids 284-380-REGION IV
US-09-134-333-13

Query Match      12.4%; Score 82; DB 9; Length 380;
Best Local Similarity 25.0%; Pred. No. 2.9;
Matches 40; Conservative 26; Mismatches 56; Indels 38; Gaps 6;

QY 8 FVATLAVMTVSGEDKKHDYQNEFDLMEIRIHEQIKKGELALFYLOEQINHFE----- 60
Db 28 YLKLPLAGMYKTIKKQLENHVNFAFNITDMLDSRLKKNRYFLEVLNSDLNPFKYSSGEY 87

QY 61 --EKPTK---EMKDKIVAEMDTIIAMI-----DGVR-----GVLDRL 92
Db 88 IIKDPYKLLDLEKKKKLIGSYKYIGASIDMDLATANDGVTYNKGELYKTHLDGVKTEI 147

QY 93 MQRKDLDFEQYNLEMLKKSGDILERDLKKEEARVKKIEV 132
Db 148 KKVED-DIKQD--BELKKLGNVNSQDSKKNEFIKKAEL 184

RESULT 7
US-09-742-096-3
; Sequence 3, Application US/09742096
; Patent No. US20020155441A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 200773USODIV
; CURRENT APPLICATION NUMBER: US/09/742,096
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/973,642
; PRIOR FILING DATE: 1998-02-06
; PRIOR APPLICATION NUMBER: PCT/FR96/00894
; PRIOR FILING DATE: 1996-06-12
; PRIOR APPLICATION NUMBER: FR 95/07007
; PRIOR FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
```

```
; ORGANISM: Homo sapiens
US-10-256-250-15

Query Match      12.2%; Score 80.5; DB 12; Length 709;
Best Local Similarity 23.6%; Pred. No. 9.1;
Matches 35; Conservative 26; Mismatches 34; Indels 53; Gaps 6;

QY 20 EDKHDYQN-----EFDLLMERIHEQIKKGELALFYLQEQ----- 55
Db 282 ETEKEDHVRNRNHNENLNAMQNLEQKFILEQREHEKLOQKELQIDSLQOEKELSSSL 341
QY 56 ---INHFEKPTKE---MKDKIVAEMDTIAMIIDGVRGVLDRMLMOR-----KDLDF 101
Db 342 HQKLCSPQEEVMVKEKNLFEELKQTLDE-----LDKLOQKEQAERLVKQLEEE 390
QY 102 EQYNLEMLKSGDILERDLKKEEARVKK 129
Db 391 AKSRAEELK----LLEEKLGKAELEK 414

RESULT 10
US-09-978-309A-47
; Sequence 47, Application US/09978309A
; Publication No. US20030100490A1
; GENERAL INFORMATION:
; APPLICANT: Cruz, Tony
; APPLICANT: Pastrak, Aleksandra
; APPLICANT: Turley, Eva A.
; TITLE OF INVENTION: Compositions and Methods for Treating Cellular Response to
; TITLE OF INVENTION: Injury and Other Proliferating Cell Disorders Regulated by
; TITLE OF INVENTION: Hyaladerin and Hyaluronans
; FILE REFERENCE: 03352-010
; CURRENT APPLICATION NUMBER: US/09/978,309A
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 09/685,010
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: US 09/541,522
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 60/127,457
; PRIOR FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 725
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-978-309A-47

Query Match      12.2%; Score 80.5; DB 11; Length 725;
Best Local Similarity 23.6%; Pred. No. 9.4;
Matches 35; Conservative 26; Mismatches 34; Indels 53; Gaps 6;

QY 20 EDKHDYQN-----EFDLLMERIHEQIKKGELALFYLQEQ----- 55
Db 298 ETEKEDHVRNRNHNENLNAMQNLEQKFILEQREHEKLOQKELQIDSLQOEKELSSSL 357
QY 56 ---INHFEKPTKE---MKDKIVAEMDTIAMIIDGVRGVLDRMLMOR-----KDLDF 101
Db 358 HQKLCSPQEEVMVKEKNLFEELKQTLDE-----LDKLOQKEQAERLVKQLEEE 406
QY 102 EQYNLEMLKSGDILERDLKKEEARVKK 129
Db 407 AKSRAEELK----LLEEKLGKAELEK 430

RESULT 11
US-10-256-250-14
; Sequence 14, Application US/10256250
; Publication No. US20030170755A1
; GENERAL INFORMATION:
; APPLICANT: Schmitt, Michael
; TITLE OF INVENTION: Tumor-Associated Antigen RHAMM
; FILE REFERENCE: L0461/70143 (JRV/MXA)
```

```
; CURRENT APPLICATION NUMBER: US/10/256,250
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/324,989
; PRIOR FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 725
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-256-250-14

Query Match      12.2%; Score 80.5; DB 12; Length 725;
Best Local Similarity 23.6%; Pred. No. 9.4;
Matches 35; Conservative 26; Mismatches 34; Indels 53; Gaps 6;

QY 20 EDKHDYQN-----EFDLLMERIHEQIKKGELALFYLQEQ----- 55
Db 298 ETEKEDHVRNRNHNENLNAMQNLEQKFILEQREHEKLOQKELQIDSLQOEKELSSSL 357
QY 56 ---INHFEKPTKE---MKDKIVAEMDTIAMIIDGVRGVLDRMLMOR-----KDLDF 101
Db 358 HQKLCSPQEEVMVKEKNLFEELKQTLDE-----LDKLOQKEQAERLVKQLEEE 406
QY 102 EQYNLEMLKSGDILERDLKKEEARVKK 129
Db 407 AKSRAEELK----LLEEKLGKAELEK 430

RESULT 12
US-10-082-830-260
; Sequence 260, Application US/10082830
; Publication No. US20030077604A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Herve
; APPLICANT: Salceda, Susana
; APPLICANT: Liu, Chenghua
; APPLICANT: Turner, Leah
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific
; TITLE OF INVENTION: Genes and Proteins
; FILE REFERENCE: DEX-0249
; CURRENT APPLICATION NUMBER: US/10/082,830
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,802
; PRIOR FILING DATE: 2000-10-27
; NUMBER OF SEQ ID NOS: 282
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 260
; LENGTH: 2383
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-082-830-260

Query Match      12.2%; Score 80.5; DB 15; Length 2383;
Best Local Similarity 22.4%; Pred. No. 42;
Matches 28; Conservative 30; Mismatches 40; Indels 27; Gaps 4;

QY 25 DYQNEFDLLMER-----IHEQIKKGELALFYLQEQINHFE---KPTKEMKDK 70
Db 355 DYQADKALTIVRSVLTTRRRQVQDLRQQLAGCQEAVALIQQHDQWEEEGKALRQRLQK 414
QY 71 IVAEMDTIAMIIDGVRGVLDRMLMORKDLDIFEQYNLEMLKSGDILERD---LKKEEARV 127
Db 415 LTGERDTLAGQTVDLQGEVDSLSKER-----ELLOKAREELRQQLVLEQEAURL 464
QY 128 KKIEV 132
Db 465 RRNVV 469

RESULT 13
US-10-171-311-164
```

```
; Sequence 164, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersh, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 164
; LENGTH: 1938
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-164
```

```
Query Match      12.1%; Score 80; DB 15; Length 1938;
Best Local Similarity 25.6%; Pred. No. 36;
Matches 33; Conservative 29; Mismatches 47; Indels 20; Gaps 7;

Qy 9 VATLAVMTVSGEDKKHDYQNEFDLLMERIHEQIKKGELALFYLOEQINHFEEKPT---- 64
Db 1366 ISTLNQLSDSKKLQDFAS-----TVEALEEGKKRFQKEIENLTQQ---YEEKAAAYDK 1417

Qy 65 -KEMDKIVAEMDTIIAMIDGVRGVLDRL--MQRK-DLDIFEQYNLEMLKSGDILERDL 120
Db 1418 LEKTKNRLQQLDLDLVLDNQRLVSNLEKKQKQKFDQLLAEBKNIS--SKYAD--ERDR 1473

Qy 121 KKEEARVKK 129
Db 1474 AEAAREKE 1482
```

```
RESULT 14
US-09-927-597-2
; Sequence 2, Application US/09927597
; Publication No. US20030032018A1
; GENERAL INFORMATION:
; APPLICANT: Malik, Fady
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; APPLICANT: Craven, Andrew
; APPLICANT: Sakowicz, Roman
; APPLICANT: Hartman, James
; TITLE OF INVENTION: Human smooth muscle myosin heavy chain
; FILE REFERENCE: CYTOP018
; CURRENT APPLICATION NUMBER: US/09/927,597
; CURRENT FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1945
; TYPE: PRT
; ORGANISM: Human
US-09-927-597-2
```

Query Match 12.1%; Score 80; DB 11; Length 1945;
Best Local Similarity 25.6%; Pred. No. 36;

```
Matches 33; Conservative 29; Mismatches 47; Indels 20; Gaps 7;

Qy 9 VATLAVMTVSGEDKKHDYQNEFDLLMERIHEQIKKGELALFYLOEQINHFEEKPT---- 64
Db 1373 ISTLNQLSDSKKLQDFAS-----TVEALEEGKKRFQKEIENLTQQ---YEEKAAAYDK 1424

Qy 65 -KEMDKIVAEMDTIIAMIDGVRGVLDRL--MQRK-DLDIFEQYNLEMLKSGDILERDL 120
Db 1425 LEKTKNRLQQLDLDLVLDNQRLVSNLEKKQKQKFDQLLAEBKNIS--SKYAD--ERDR 1480

Qy 121 KKEEARVKK 129
Db 1481 AEAAREKE 1489
```

```
RESULT 15
US-10-171-311-162
; Sequence 162, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersh, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 162
; LENGTH: 1972
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-162
```

```
Query Match      12.1%; Score 80; DB 15; Length 1972;
Best Local Similarity 25.6%; Pred. No. 37;
Matches 33; Conservative 29; Mismatches 47; Indels 20; Gaps 7;

Qy 9 VATLAVMTVSGEDKKHDYQNEFDLLMERIHEQIKKGELALFYLOEQINHFEEKPT---- 64
Db 1366 ISTLNQLSDSKKLQDFAS-----TVEALEEGKKRFQKEIENLTQQ---YEEKAAAYDK 1417

Qy 65 -KEMDKIVAEMDTIIAMIDGVRGVLDRL--MQRK-DLDIFEQYNLEMLKSGDILERDL 120
Db 1418 LEKTKNRLQQLDLDLVLDNQRLVSNLEKKQKQKFDQLLAEBKNIS--SKYAD--ERDR 1473

Qy 121 KKEEARVKK 129
Db 1474 AEAAREKE 1482
```

Search completed: November 17, 2003, 08:29:58
Job time : 30 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 17, 2003, 08:22:19 ; Search time 22 Seconds
(without alignments)
253.865 Million cell updates/sec

Title: US-09-877-160A-1
Perfect score: 662
Sequence: 1 MKFIIAFVATLAVMTVSGE.....GDILDLKKEARVKKIEV 132

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep: *
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep: *
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep: *
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep: *
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep: *
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	82.5	12.5	282	4 US-09-107-532A-3671	Sequence 3671, Ap
2	82.5	12.5	342	4 US-09-107-532A-5664	Sequence 5664, Ap
3	82	12.4	1786	3 US-08-973-462-8	Sequence 8, Appli
4	81.5	12.3	344	6 5210183-2	Patent No. 5210183
5	81.5	12.3	683	6 5210183-3	Patent No. 5210183
6	80	12.1	816	2 US-08-533-306A-6	Sequence 6, Appli
7	80	12.1	816	2 US-08-742-923A-6	Sequence 6, Appli
8	80	12.1	885	2 US-08-533-306A-4	Sequence 4, Appli
9	80	12.1	885	2 US-08-742-923A-4	Sequence 4, Appli
10	79	11.9	306	4 US-09-107-532A-4841	Sequence 4841, Ap
11	79	11.9	319	4 US-09-107-532A-4723	Sequence 4723, Ap
12	79	11.9	413	4 US-09-198-452A-229	Sequence 229, App
13	78	11.8	142	4 US-09-328-352-6620	Sequence 6620, Ap
14	78	11.8	984	1 US-08-242-932-2	Sequence 2, Appli
15	78	11.8	984	1 US-08-714-481-2	Sequence 2, Appli
16	78	11.8	984	5 PCT-US95-06111-2	Sequence 2, Appli
17	77.5	11.7	475	4 US-09-370-838-193	Sequence 193, App
18	76.5	11.6	107	1 US-08-182-175A-105	Sequence 105, App
19	76.5	11.6	107	1 US-08-474-633A-92	Sequence 92, Appl
20	76.5	11.6	107	4 US-08-823-771-92	Sequence 92, Appl
21	76.5	11.6	107	5 PCT-US92-06412-105	Sequence 105, App
22	76	11.5	187	4 US-09-328-352-5117	Sequence 5117, Ap
23	76	11.5	431	4 US-09-286-981B-3	Sequence 3, Appli
24	75	11.3	1010	4 US-09-134-001C-5178	Sequence 5178, Ap
25	74	11.2	281	4 US-09-914-259-63	Sequence 63, Appl
26	74	11.2	885	4 US-09-107-532A-5104	Sequence 5104, Ap
27	73.5	11.1	251	4 US-09-286-981B-4	Sequence 4, Appli

28	73.5	11.1	414	4 US-09-286-981B-10	Sequence 10, Appl
29	73.5	11.1	496	3 US-08-740-223A-16	Sequence 16, Appl
30	73.5	11.1	496	4 US-09-709-188-16	Sequence 16, Appl
31	73.5	11.1	564	3 US-09-308-022-6	Sequence 6, Appli
32	73.5	11.1	631	3 US-08-847-065-25	Sequence 25, Appl
33	73.5	11.1	835	2 US-08-968-751-4	Sequence 4, Appli
34	73	11.0	258	4 US-09-198-452A-563	Sequence 563, App
35	73	11.0	284	4 US-09-167-206-10	Sequence 10, Appl
36	73	11.0	284	4 US-09-914-259-40	Sequence 40, Appl
37	73	11.0	284	4 US-09-914-259-48	Sequence 48, Appl
38	73	11.0	284	4 US-09-914-259-50	Sequence 50, Appl
39	73	11.0	285	1 US-08-149-809-24	Sequence 24, Appl
40	72	10.9	388	2 US-08-705-868-4	Sequence 4, Appli
41	72	10.9	388	3 US-09-123-615-4	Sequence 4, Appli
42	72	10.9	421	4 US-09-585-173B-52	Sequence 52, Appl
43	72	10.9	1326	4 US-09-252-991A-17932	Sequence 17932, A
44	72	10.9	2482	1 US-08-328-254-6	Sequence 6, Appli
45	72	10.9	3248	1 US-08-353-700-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-107-532A-3671
; Sequence 3671, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3671:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 282 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...282
; SEQUENCE DESCRIPTION: SEQ ID NO: 3671:

US-09-107-532A-3671

Query Match 12.5%; Score 82.5; DB 4; Length 282;
Best Local Similarity 21.2%; Pred. No. 0.37;
Matches 28; Conservative 34; Mismatches 33; Indels 37; Gaps 6;

Qy 23 KHDYQNEFDLLMERIHEQIKKGELALFYLOEQINHFEEK---PTKEMKDKIVAEMDTII 79
Db 89 KHDYQN-----LLFSLHSYISEGDL-----EGLRSYYDQILPTROMNLFPA-----L 133

Qy 80 AMIDGV-----RGVLDRLMQRKOLDIFEQYNLEMLKKSGLILERDLK 121
Db 134 SMLDNMKVPEIRSLSLKLMMAQEKGLIVQHFPEPIELDTKHTVNLVRLGILDNAIE 193

Qy 122 -KEEARVKKIEV 132
Db 194 GASAAXKKKIEL 205

RESULT 2
US-09-107-532A-5664
; Sequence 5664, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5664:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 342 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...342
; SEQUENCE DESCRIPTION: SEQ ID NO: 5664:
US-09-107-532A-5664

Query Match 12.5%; Score 82.5; DB 4; Length 342;
Best Local Similarity 28.4%; Pred. No. 0.47;
Matches 29; Conservative 18; Mismatches 48; Indels 7; Gaps 2;

Qy 36 ERIHEQIKKG---ELALFYLOEQINHFEEKPTKEMKDKIVA---EMDTIIAMIDGVRGV 88
Db 38 ERILDLVKKGILSTEGDLLESMAATEKDEKQIKKADKVTASHKEKDQASQLDKLENG 97

Qy 89 LDRLMQRKOLDIFEQYNLEMLKKSGLILERDLKKEEARVKKI 130
Db 98 EEEISEPVDPKERERQDQENLEKILDELATEANKTSARLDEV 139

RESULT 3
US-08-973-462-8
; Sequence 8, Application US/08973462B
; Patent No. 6191270
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 0660-0125-0 PCT
; CURRENT APPLICATION NUMBER: US/08/973,462B
; CURRENT FILING DATE: 1998-02-06
; EARLIER APPLICATION NUMBER: PCT/FR96/00894
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: FR 95/07007
; EARLIER FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Polypeptide
US-08-973-462-8

Query Match 12.4%; Score 82; DB 3; Length 1786;
Best Local Similarity 24.8%; Pred. No. 4.5;
Matches 27; Conservative 29; Mismatches 41; Indels 12; Gaps 4;

Qy 28 NEFDLLMERIHEQIKKGELALFYLOEQINHFEEKPTKEMKDKIVAEMDTIIAMID---G 84
Db 1237 NDTSKLIEETQE-LNEVEADLIKDMEKLEKALSSEDSKEIIDAKDDTLEKVEEHD 1295

Qy 85 VRGVLDRLMQRKOLDIFEQYNLEMLKKSGLI---LERDLKKEEARVKKIE 131
Db 1296 ITTTLDEVVELKDVE-----EDKIEKVSCLKLEEDILKEVKEIKELE 1338

RESULT 4
5210183-2
; Patent No. 5210183
; APPLICANT: LINDAHL, GUNNAR; FRITHZ, ELISABET; HEDEN, LARS-OLOF
; TITLE OF INVENTION: PROTEIN ARP, WITH IMMUNOGLOBULIN A
; BINDING ACTIVITY, THE CORRESPONDING VECTORS AND HOSTS, REAGENT
; KIT AND PHARMACEUTICAL COMPOSITION
; NUMBER OF SEQUENCES: 3
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/270,099
; FILING DATE: 14-NOV-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 186,097
; FILING DATE: 25-APR-1988
; SEQ ID NO: 2
; LENGTH: 344
5210183-2

Query Match 12.3%; Score 81.5; DB 6; Length 344;
Best Local Similarity 25.8%; Pred. No. 0.61;
Matches 32; Conservative 27; Mismatches 50; Indels 15; Gaps 5;

Qy 22 KKHQYQNEFDLLMERIHEQIKKGELALFYLOEQINHFEEKPTKEMKD--KIVAEMDTII 79
Db 102 REGQYQDKIELEKERKEKERQERQELERQYQIEADKHQYEQQKQHQEQQLEAEKQKLA 161

Query Match 12.1%; Score 80; DB 2; Length 816;
Best Local Similarity 25.6%; Pred. No. 2.7;
Matches 33; Conservative 29; Mismatches 47; Indels 20; Gaps 7;
QY 9 VATLAVMTVSGEDKKHDYQNEFDLLMERIHEQIKKGELALFYLQEQINHFEEKPT---- 64
Db 244 ISTLNQLSDSKKLQDFAS-----TVEALEEGKKRFQKEIENLTQQ---YEEKAAAYDK 295
QY 65 -KEMKDKIVAEMDTIAMIIDGVRGVLDR--MQRK-DLDIFEQYNLEMLKKSGLERDL 120
Db 296 LEKTKNRLQQLDLDLVVDLDNQRLVSNLEKKQKQKFDQLLAEEKNIS--SKYAD--ERDR 351
QY 121 KKEEARVKK 129
Db 352 AEAAREKE 360

RESULT 8
US-08-533-306A-4
; Sequence 4, Application US/085333306A
; Patent No. 5837457
; GENERAL INFORMATION:
; APPLICANT: Liu, Pu
; APPLICANT: Collins, Francis S.
; APPLICANT: Siciliano, Michael J.
; APPLICANT: Claxton, David
; TITLE OF INVENTION: Markers for Detection of Chromosome 16
; TITLE OF INVENTION: Rearrangements
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
; STREET: P.O. Box 828
; CITY: Bloomfield Hills
; STATE: MI
; COUNTRY: USA
; ZIP: 48303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/533,306A
; FILING DATE: September 25, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, DeAnn F.
; REGISTRATION NUMBER: 36683
; REFERENCE/DOCKET NUMBER: 2115-00869COB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 641-1600
; TELEFAX: (810) 641-0270
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 885 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-533-306A-4

Query Match 12.1%; Score 80; DB 2; Length 885;
Best Local Similarity 25.6%; Pred. No. 3;
Matches 33; Conservative 29; Mismatches 47; Indels 20; Gaps 7;
QY 9 VATLAVMTVSGEDKKHDYQNEFDLLMERIHEQIKKGELALFYLQEQINHFEEKPT---- 64
Db 313 ISTLNQLSDSKKLQDFAS-----TVEALEEGKKRFQKEIENLTQQ---YEEKAAAYDK 364
QY 65 -KEMKDKIVAEMDTIAMIIDGVRGVLDR--MQRK-DLDIFEQYNLEMLKKSGLERDL 120
Db 365 LEKTKNRLQQLDLDLVVDLDNQRLVSNLEKKQKQKFDQLLAEEKNIS--SKYAD--ERDR 420

QY 121 KKEEARVKK 129
Db 421 AEAAREKE 429
RESULT 9
US-08-742-923A-4
; Sequence 4, Application US/08742923A
; Patent No. 5869611
; GENERAL INFORMATION:
; APPLICANT: Liu, Pu
; APPLICANT: Collins, Francis S.
; APPLICANT: Siciliano, Michael J.
; APPLICANT: Claxton, David
; TITLE OF INVENTION: Markers for Detection of Chromosome 16
; TITLE OF INVENTION: Rearrangements
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
; STREET: P.O. Box 828
; CITY: Bloomfield Hills
; STATE: MI
; COUNTRY: USA
; ZIP: 48303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,923A
; FILING DATE: No. 5869611ember 1, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, DeAnn F.
; REGISTRATION NUMBER: 36683
; REFERENCE/DOCKET NUMBER: 2115-00869DVC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 641-1600
; TELEFAX: (810) 641-0270
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 885 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-742-923A-4

Query Match 12.1%; Score 80; DB 2; Length 885;
Best Local Similarity 25.6%; Pred. No. 3;
Matches 33; Conservative 29; Mismatches 47; Indels 20; Gaps 7;
QY 9 VATLAVMTVSGEDKKHDYQNEFDLLMERIHEQIKKGELALFYLQEQINHFEEKPT---- 64
Db 313 ISTLNQLSDSKKLQDFAS-----TVEALEEGKKRFQKEIENLTQQ---YEEKAAAYDK 364
QY 65 -KEMKDKIVAEMDTIAMIIDGVRGVLDR--MQRK-DLDIFEQYNLEMLKKSGLERDL 120
Db 365 LEKTKNRLQQLDLDLVVDLDNQRLVSNLEKKQKQKFDQLLAEEKNIS--SKYAD--ERDR 420

RESULT 10
US-09-107-532A-4841
; Sequence 4841, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

;/ NUMBER OF SEQUENCES: 7310
;/ CORRESPONDENCE ADDRESS:
;/ ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;/ STREET: 100 Beaver Street
;/ CITY: Waltham
;/ STATE: Massachusetts
;/ COUNTRY: USA
;/ ZIP: 02354
;/
;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: CD/ROM ISO9660
;/ COMPUTER: PC
;/ OPERATING SYSTEM: <Unknown>
;/ SOFTWARE: ASCII
;/
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/09/107,532A
;/ FILING DATE: 30-Jun-1998
;/
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: 60/085,598
;/ FILING DATE: 14 May 1998
;/ APPLICATION NUMBER: 60/051571
;/ FILING DATE: July 2, 1997
;/
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: Ariniello, Pamela Deneke
;/ REGISTRATION NUMBER: 40,489
;/ REFERENCE/DOCKET NUMBER: GTC-012
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: (781)893-5007
;/ TELEFAX: (781)893-8277
;/
;/ INFORMATION FOR SEQ ID NO: 4841:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 306 amino acids
;/ TYPE: amino acid
;/ TOPOLOGY: linear
;/ MOLECULE TYPE: protein
;/ HYPOTHETICAL: YES
;/ ORIGINAL SOURCE:
;/ ORGANISM: Enterococcus faecium
;/
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (B) LOCATION 1...306
;/
;/ SEQUENCE DESCRIPTION: SEQ ID NO: 4841:
;/
US-09-107-532A-4841

Query Match 11.9%; Score 79; DB 4; Length 306;
Best Local Similarity 23.1%; Pred. No. 0.97;
Matches 30; Conservative 30; Mismatches 48; Indels 22; Gaps 5;

QY 9 VATLAVMTVSGEDKKHDYQNEFDLLMERIHEQIKKGELALFYLQEQINHFEKPTKEMK 68
Db 94 IDSLITQLTDGIQRYGSQNG---VSISRMAER-----FALDENFVTVEEKPEESME 142

QY 69 DKIVAEMDTIAMIIDGVRGVLDLDR-----LMQRKDLDFEQYNLEMLKSGDILERDLKK 122
Db 143 ELAQLALDTLKSALDSIEASREKEGQALALVKKNTTEFK---EVLKLDLSFVEIYEKE 198

QY 123 EEARV-KKIE 131
Db 199 FQIRYQKKLE 208

RESULT 11
US-09-107-532A-4723
; Sequence 4723, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham

;/ STATE: Massachusetts
;/ COUNTRY: USA
;/ ZIP: 02354
;/
;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: CD/ROM ISO9660
;/ COMPUTER: PC
;/ OPERATING SYSTEM: <Unknown>
;/ SOFTWARE: ASCII
;/
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/09/107,532A
;/ FILING DATE: 30-Jun-1998
;/
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: 60/085,598
;/ FILING DATE: 14 May 1998
;/ APPLICATION NUMBER: 60/051571
;/ FILING DATE: July 2, 1997
;/
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: Ariniello, Pamela Deneke
;/ REGISTRATION NUMBER: 40,489
;/ REFERENCE/DOCKET NUMBER: GTC-012
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: (781)893-5007
;/ TELEFAX: (781)893-8277
;/
;/ INFORMATION FOR SEQ ID NO: 4723:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 319 amino acids
;/ TYPE: amino acid
;/ TOPOLOGY: linear
;/ MOLECULE TYPE: protein
;/ HYPOTHETICAL: YES
;/ ORIGINAL SOURCE:
;/ ORGANISM: Enterococcus faecium
;/
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (B) LOCATION 1...319
;/
;/ SEQUENCE DESCRIPTION: SEQ ID NO: 4723:
;/
US-09-107-532A-4723

Query Match 11.9%; Score 79; DB 4; Length 319;
Best Local Similarity 20.7%; Pred. No. 1;
Matches 29; Conservative 32; Mismatches 43; Indels 36; Gaps 5;

QY 8 FVATLAVMTVSGEDKKHDYQNEFDLLMERIHEQIKKGELALF--YLQEQINHFEKPTK 65
Db 164 FYSTIFEMT---ETEHTVRDESLLKIDEQMSKQVRESELEAFKQEAEEQIQLNVNTPAQ 220

QY 66 EMKDKIVAEMDTIAMIIDGVRGVLDRLMQRKDLDFEQYNLEMLKKS-----G 113
Db 221 QQMTRYLVDDQGIIVN-----DVLNEKKTEELKQSAREAVQPMVIFQG 262

QY 114 DILERDLKKEEA-RVKKIEV 132
Db 263 EIIVRESNQIDASAMKKLEL 282

RESULT 12
US-09-198-452A-229
; Sequence 229, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Grifais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragme
; thereof and uses thereof, in particular for the diagnosis, pr
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 229
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:

SEQUENCE CHARACTERISTICS:

LENGTH: 984 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-714-481-2

Query Match 11.8%; Score 78; DB 1; Length 984;
Best Local Similarity 17.5%; Pred. No. 5.6;
Matches 30; Conservative 37; Mismatches 50; Indels 54; Gaps 4;

QY	16	TVSGEDKKHDYQ-----NEFDLLMERIHEQIKKGELA-----	48
		: : : : :	
Db	88	TILSHEQKNEFKTKIDETNDSALLELENQFNETNRLHLIKQHEEVEKDKKAKQQTILKQ	147
		: : : : :	
QY	49	-----LPYLQEQINHEEKPTKEMKDIVA--EMDTIAMIIDGVRGVLDRLMQRKDLDI	100
		: : : : :	
Db	148	SDTKVDLSNIDKELNHQKSQVEKMAEQKGTNEDKDSMLKKIEDIRKQAQQAADKKEDAEV	207
		: : : : :	
QY	101	FEQYNLEMLKKSGDILE-----RDLKKEEARVKKIEV	132
		: : : : :	
Db	208	KVQLEEEAHSKLKQVVEDFRKKFKTSEQVTPKKRVKRDLAANENNQQKIEL	258
		: : : : :	

Search completed: November 17, 2003, 08:25:34
Job time : 23 secs